

**Bewertung der diagnostischen Information
morphologischer und kinetischer Kriterien in der
dynamischen MR-Mammographie**

**Dissertation
zur Erlangung des akademischen Grades
doctor medicinae (Dr. med.)**

vorgelegt dem Rat der Medizinischen Fakultät
der Friedrich-Schiller-Universität Jena

von Tobias Gröschel
geboren am 13. August 1982 in Jena

Gutachter:

1. Prof. Dr. med. Dipl.-Chem. W.A. Kaiser, Friedrich-Schiller-Universität Jena
2. Prof. Dr. med. Ingo B. Runnebaum, MBA, Friedrich-Schiller-Universität Jena
3. Prof. Dr. med. Rüdiger Schulz-Wendtland, Friedrich-Alexander-Universität Erlangen

Tag der öffentlichen Verteidigung: 04.07.2011

Abkürzungsverzeichnis

ACR:	American College of Radiology
ADC:	apparent diffusion coefficient
ADH:	atypische duktale Hyperplasie
AUC:	area under curve
BI-RADS:	breast imaging reporting and data system
BRCA:	breast cancer susceptibility gene
CAD:	computer aided diagnosis
CUP:	carcinoma of unknown primary
diff. ips.:	diffus ipsilateral
DCIS:	duktales Carcinoma in situ
EUSOBI:	European Society of Breast Imaging
FFE:	fast field echo
FLASH:	fast low angle shot
FNB:	Feinnadelaspirationszytologie
FSU:	Friedrich-Schiller-Universität
Gd-DTPA:	Gadolinium-Diethylentriaminpentaacetat
HRT:	hormone replacement therapy (Hormonersatztherapie)
IDC:	invasives duktales Karzinom
ILC:	invasives lobuläres Karzinom
IDIR:	Institut für Diagnostische und Interventionelle Radiologie
IOR:	Interobserverreliabilität
KG:	Körpergewicht
KI:	Konfidenzintervall
KM:	Kontrastmittel
LCIS:	lobuläres Carcinoma in situ
ln:	natürlicher Logarithmus
MR:	Magnetresonanz
MRM:	Magnetresonanzmammographie
MRT:	Magnetresonanztomographie
NPV:	negativer prädiktiver Wert
OR:	Odds Ratio
PPV:	positiver prädiktiver Wert

RKI:	Robert-Koch-Institut
ROC:	receiver operating characteristic
ROI:	region of interest
SE:	spin echo
T:	Tesla
T1:	longitudinale Relaxationszeit
T2:	transversale Relaxationszeit
T1w:	T1-gewichtet
T2w:	T2-gewichtet
TE:	Echozeit
TR:	Relaxationszeit
TSE:	turbo spin echo
UKJ:	Universitätsklinikum Jena
VEGF:	vascular endothelial growth factor
Z. n.:	Zustand nach

Inhaltsverzeichnis

Inhaltsverzeichnis	i
Tabellenverzeichnis	iv
Abbildungsverzeichnis	v
1 Zusammenfassung	1
2 Einleitung	3
2.1 Epidemiologie des Mammakarzinoms	3
2.2 Veränderungen der Mamma	6
2.2.1 Bösartige Veränderungen	6
2.2.2 Gutartige Veränderungen	6
2.3 Mammadiagnostik	7
2.3.1 Klinische Untersuchung	7
2.3.2 Mammographie	7
2.3.3 Sonographie	8
2.3.4 MR-Mammographie	9
2.3.5 Invasive Verfahren zur histologischen Diagnostik	12
2.4 Beschreibung MR-mammographisch sichtbarer Läsionen	12
2.5 MR-mammographische Interpretationsmodelle	15
3 Ziele dieser Arbeit	16
4 Methoden	18
4.1 Einschluss-/ Ausschlusskriterien	18
4.2 Einteilung der histopathologischen Befunde	19

4.3	Untersuchungsablauf	21
4.4	Auswertung und Datenerfassung.....	26
4.4.1	Beurteilung der morphologischen Kriterien.....	26
4.4.2	Beurteilung der dynamischen Kriterien	28
4.4.3	Beurteilung des die Läsion umgebenden Gewebes.....	31
4.5	Histologische Aufarbeitung	32
4.6	Statistische Methoden	33
5	Ergebnisse	35
5.1	Häufigkeit von MR-Zeichen.....	35
5.2	Binäre logistische Regressionsmodelle	39
5.2.1	Einschlussmethode, Vorwärts- und Rückwärtsselektion	39
5.2.2	Binäre logistische Modelle mit fünf Deskriptoren	40
5.3	Erfassung von Läsionen mit sicherer Klassifizierbarkeit	44
5.3.1	Erfassung sicher maligner Läsionen	44
5.3.2	Erfassung sicher benigner Läsionen.....	47
5.3.3	Summarische Erfassung sicher maligner und benigner Läsionen.....	48
5.3.4	Zeichenkombinationen bei sicher klassifizierbaren Läsionen	49
6	Diskussion.....	52
6.1	Interpretationsmodelle in der MRM-Diagnostik	52
6.2	Methodische Aspekte	55
6.3	Diskussion der für die Regressionsmodelle verwendeten Deskriptoren	59
6.4	Vergleich mit anderen publizierten MRM-Interpretationsmodellen.....	72
6.5	Publikationsvergleich hinsichtlich einer sicheren Dignitätszuweisung.....	75
6.6	Ausblick.....	77

7	Schlussfolgerungen	79
8	Literaturverzeichnis	81
9	Anhang.....	98
9.1	Beispiel und Formelherleitungen	98
9.1.1	Berechnung der Malignitätswahrscheinlichkeit	98
9.1.2	Formelherleitung	100
9.2	Danksagung.....	102
9.3	Curriculum vitae	103
9.4	Ehrenwörtliche Erklärung	104

Tabellenverzeichnis

Tabelle 2.1:	BI-RADS-Einstufung des ACR	14
Tabelle 4.1:	Häufigkeiten der histologischen Untergruppen gutartiger Befunde	19
Tabelle 4.2:	Häufigkeiten der histologischen Untergruppen maligner Befunde.....	20
Tabelle 4.3:	Anzahl der von den unterschiedlichen Tomographen erfassten Läsionen	22
Tabelle 4.4:	Messprotokoll der MR-Mammographie mit den Philips-Geräten	23
Tabelle 4.5:	Messprotokoll der MR-Mammographie mit dem Siemens Vision	24
Tabelle 4.6:	Messprotokoll der MR-Mammographie mit dem Siemens Symphony.....	25
Tabelle 4.7:	Deskriptoren der Läsionsmorphologie	28
Tabelle 4.8:	Beurteilungsmöglichkeiten für dynamische Deskriptoren.....	31
Tabelle 4.9:	Deskriptoren der die Läsion umgebenden Auffälligkeiten	32
Tabelle 5.1:	Häufigkeiten morphologischer Zeichen	36
Tabelle 5.2:	Häufigkeiten KM-dynamischer Zeichen	37
Tabelle 5.3:	Häufigkeiten der die Läsionsumgebung beschreibenden Zeichen	38
Tabelle 5.4:	Statistische Kennzahlen der initial errechneten Regressionsmodelle	40
Tabelle 5.5:	AUC-Vergleich der initial errechneten Modelle	40
Tabelle 5.6:	Statistische Kennwerte zu den Modellen mit fünf Deskriptoren	41
Tabelle 5.7:	AUC-Vergleich initial Erstellter mit den Fünf-Deskriptoren-Modellen.....	41
Tabelle 5.8:	Optimiertes fünf Deskriptoren enthaltendes Regressionsmodell	42
Tabelle 5.9:	Fünf-Deskriptoren-Regressionsmodell mit konventionellen Zeichen	43
Tabelle 5.10:	Modellabhängige Erfassung sicher maligner Befunde.....	45
Tabelle 5.11:	Modellabhängige Erfassung sicher benigner Befunde.....	47
Tabelle 5.12:	Modellabhängige Erfassung sicher maligner und benigner Befunde.....	49
Tabelle 5.13:	Mögliche Merkmalskombinationen für sicher maligne Befunde.....	50
Tabelle 5.14:	Mögliche Merkmalskombinationen für sicher benigne Befunde.....	51
Tabelle 9.1:	Deskriptorvariablen für die Berechnung der Malignitätswahrscheinlichkeit...	99

Abbildungsverzeichnis

Abbildung 2.1:	Prävalenz und Inzidenz der zehn häufigsten Krebsformen.....	4
Abbildung 4.1:	Root- und Hook-Zeichen.....	27
Abbildung 4.2:	Schematische Darstellung KM-dynamischer Zeichen	29
Abbildung 4.3:	Schematische Darstellung des zeitlichen KM-Anreicherungsverhaltens	30
Abbildung 4.4:	Perifokales Ödem und zuführendes Gefäß.....	32
Abbildung 5.1:	Sicher maligner Befund.....	46
Abbildung 5.2:	Sicher gutartiger Befund	48
Abbildung 6.1:	Non-mass-Läsion und Herdbefund	62
Abbildung 6.2:	Zentripetales Enhancement	64
Abbildung 6.3:	Perifokales Ödem	67
Abbildung 6.4:	Root-sign	70

1 Zusammenfassung

Brustkrebs ist die sowohl weltweit als auch in der BRD häufigste maligne Erkrankung des weiblichen Geschlechts. Jede zehnte Frau in Deutschland erkrankt hieran im Laufe ihres Lebens. Diese epidemiologische Tatsache unterstreicht die Notwendigkeit einer frühzeitigen und treffsicheren Diagnostik, um eine schnellstmögliche Therapieeinleitung zu gewährleisten. Mit der Magnetresonanzmammographie (MRM) existiert ein hochoempfindliches Verfahren zur Detektion suspekter senologischer Veränderungen. Trotz der Möglichkeit, fast jede Krebserkrankung durch dieses bildgebende Verfahren zu erkennen, müssen falsch-positiv Befunde nach Möglichkeit vermieden werden, um unnötige Kosten sowie Belastungen für die betroffenen Patientinnen zu beschränken.

Ziel dieser Arbeit war die Analyse des diagnostischen Nutzens morphologischer und dynamischer Deskriptoren zur Differentialdiagnose in der MRM. Anhand einer großen Fallzahl sollte die Genauigkeit von Deskriptorkombinationen anhand statistischer Modelle beurteilt werden. Diese sollten für den klinischen Einsatz auf möglichst wenigen, aussagekräftigen Deskriptoren beruhen, ohne dabei die Treffsicherheit der Methode maßgeblich zu verschlechtern. Zuletzt sollte auch geprüft werden, ob Subgruppen mit an 100% grenzender Wahrscheinlichkeit klassifiziert werden können.

Zwischen dem 01.12.1994 und dem 16.10.2006 wurden 5956 MR-Mammographien am IDIR (Institut für Diagnostische und Interventionelle Radiologie) durchgeführt. Aus diesen Untersuchungen gelang die Rekrutierung von 1012 Patientinnen mit 1084 Kontrastmittel (KM) anreichernden histologisch aufgearbeiteten Läsionen ohne vorausgegangene Intervention. Die KM-verstärkte MRM erfolgte an 1,5-Tesla-MR-Tomographen nach einem standardisierten Untersuchungsprotokoll (T1-gewichtete dynamische KM-verstärkte sowie T2-gewichtete Sequenzen). Die Untersuchungsergebnisse wurden durch in der MRM erfahrene Radiologen im Konsens beurteilt und in einer Datenbank erfasst. Es folgte der Vergleich mit den histologischen Befunden. Das standardisierte Auswertungsprotokoll umfasste morphologische (Form, Rand, Septen, T1-Signalintensität, T1-Binnenstruktur, T2-Signalintensität, Root- und Hook-sign) und KM-dynamische (räumliches und zeitliches KM-Anreicherungsverhalten, Blooming-sign) Merkmale sowie die Beurteilung des peripher der Läsion befindlichen Gewebes (Ödem, abnorme Gefäßzeichnung, Cutisverdickung, Mamillensaumunterbrechung).

Mittels Chi-Quadrat-Test oder dem exakten Test nach Fisher wurde auf unterschiedliche Verteilung einzelner Merkmalsausprägungen zwischen benignen und malignen Veränderungen untersucht. Unter Verwendung von binärer logistischer Regressionsanalyse wurden beurteilte Kriterien in Modellen zusammengefasst und die Wahrscheinlichkeit für das Vorliegen eines bösartigen Prozesses berechnet.

Unter Einschluss aller Deskriptoren gelang eine zu 83,8% korrekte Klassifizierung suspekter Läsionen. Bei Beschreibung von lediglich fünf Merkmalen (zeitliches und räumliches KM-Anreicherungsverhalten, Root-sign, Ödem, T2-Signalintensität) verringerte sich die Genauigkeit geringfügig auf 81,5%, wobei der Unterschied signifikant war ($P=0,001$). Ein alternatives Modell mit fünf „konventionellen“ zahlreich publizierten Deskriptoren (Form, Rand, Septen, zeitl./räuml. KM-Anreicherungsmuster) klassifizierte Herdbefunde zu 80,4% korrekt, unterschied sich jedoch signifikant zu beiden vorbeschriebenen Modellen ($P<0,001$ bzw. $P=0,017$).

Hinsichtlich der sicheren Differenzierung zwischen Malignom und benigner Veränderung ergaben sich modellspezifische Vor- und Nachteile. Unter Einschluss sämtlicher evaluierter Zeichen konnte der höchste Anteil aller Läsionen (27,6%) mit großer Sicherheit ($>95\%$) richtig klassifiziert werden. Das vereinfachte mit den oben aufgeführten fünf Deskriptoren optimierte Modell ermöglichte die sichere Klassifikation von 26,9% der Malignome. Bei Betrachtung aller Zeichen war dieser Anteil auf 35,6% steigerbar. Die Fehlerquoten waren beim vereinfachten Modell jedoch etwas geringer (1,1% vs. 1,7%). Vorteil des Modells mit den etablierten Deskriptorvariablen war die Extraktion sicher gutartiger Läsionen bei noch moderater Fehlerquote (2,9%), welche zu 15,4% gelang. Die anderen Modelle lagen hier im Gegenzug in ca. 7% falsch. Die in dieser Arbeit berichteten diagnostischen Genauigkeiten beschreiben nicht die Treffsicherheit der Methode als solche, sondern lediglich mathematische Modelle subjektiver Deskriptoren in dem hier definierten Patientengut.

Insgesamt kann unter Nutzung der hier genutzten MR-Zeichen die Dignität eines Teils der untersuchten Läsionen sicher vorhergesagt werden. Die neuen Deskriptoren bedürfen jedoch weiterer prospektiver Analysen hinsichtlich ihrer Reliabilität und Validität. In der Literatur wird zunehmend zwischen soliden Herdbefunden und nicht soliden Infiltrationen (oder non-mass-Läsionen) unterschieden. Alternative Modelle könnten diese Trennung adressieren. Eine Erhöhung der diagnostischen Genauigkeit ist durch Berücksichtigung anatomischer Wachstumsmuster potenziell denkbar.

2 Einleitung

2.1 Epidemiologie des Mammakarzinoms

Das Mammakarzinom ist die häufigste Krebserkrankung der Frau in den westlichen Industrienationen sowie in der Bundesrepublik Deutschland (Bray et al. 2002). Wie aus Abb. 2.1 ersichtlich wird, ergibt sich bei weltweit 1,15 Mio. Erkrankten jährlich dieselbe Relevanz, zudem besitzt Brustkrebs die höchste Prävalenz unter allen malignen Erkrankungen (Parkin et al. 2005).

Gemäß dem vom Robert-Koch-Institut (RKI) herausgegebenen Bericht „Krebs in Deutschland 2003-2004“ erkrankten hierzulande 57.000 Frauen jährlich an Brustkrebs, was 27,8% aller hiesigen Krebserkrankungsfälle der Frau entspricht. Das Risiko einer Frau, bis zum 80. Lebensjahr an Brustkrebs zu erkranken, wird mit 10% angegeben. Mit 63 Jahren liegt das mittlere Erkrankungsalter 6 Jahre unter dem mittleren Erkrankungsalter aller Krebserkrankungen (Batzler et al. 2008). Das Auftreten des Mammakarzinoms ist jedoch nicht ausschließlich auf das höhere Lebensalter beschränkt. In der Gruppe der Frauen unter 60 Jahren beträgt der Anteil des Brustkrebses an allen Krebserkrankungen 40% (Batzler 2006). Weiterhin berichtet das RKI von einer stetigen Zunahme der Inzidenz in Deutschland seit 1980, während die Mortalität seit Mitte der 1990er Jahre leicht sinken soll. Das Krebsregister Saarland kam ebenfalls zu dem Ergebnis einer leicht fallenden Mortalitätsrate seit Mitte der 1990er Jahre. Im Jahr 2000 hatte Brustkrebs einen Anteil von 4,5% sämtlicher Todesursachen beim weiblichen Geschlecht im Saarland (Krebsregister Saarland).

Trotz der eminenten Bedeutung des Mammakarzinoms innerhalb der gynäkologischen Onkologie darf nicht vergessen werden, dass auch Männer betroffen sein können. Sie stellen ca. 1% aller an Brustkrebs Erkrankten (Lanitis et al. 2008).

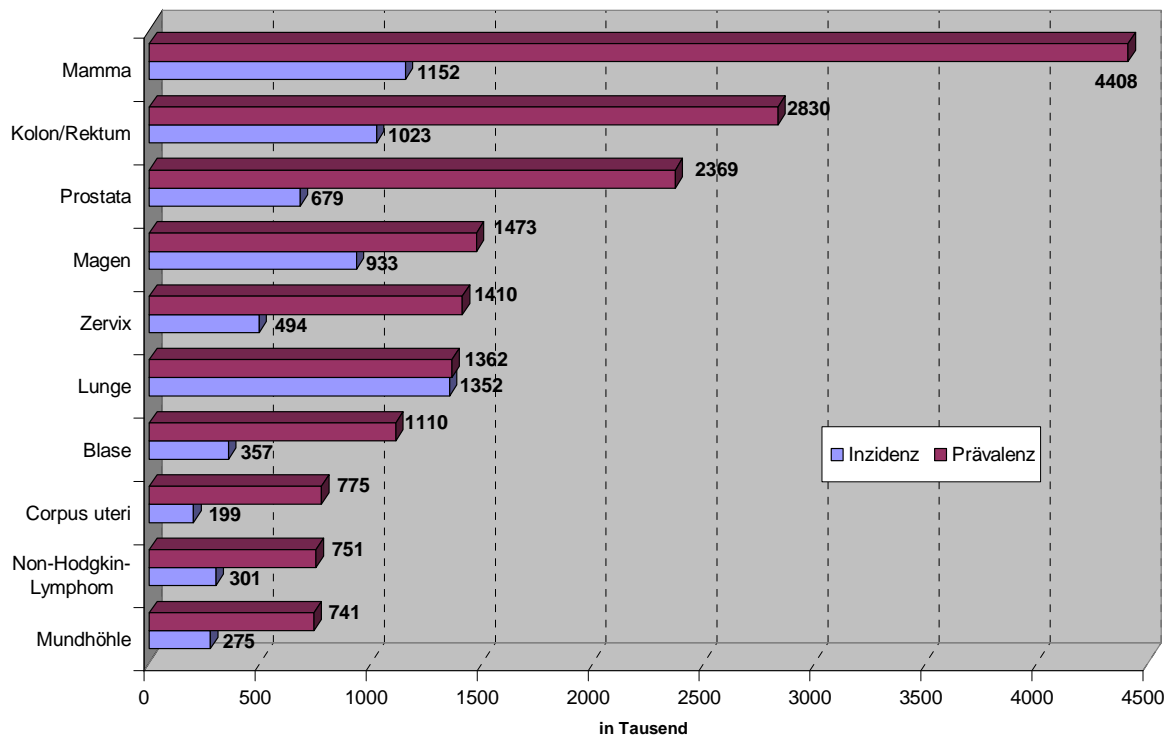


Abbildung 2.1: Prävalenz und Inzidenz der zehn häufigsten Krebsformen

Erfassung der weltweit zehn häufigsten Krebsformen. Mit ca. 4,4 Mio. rangiert der Brustkrebs vor dem Dickdarmkrebs (2,8 Mio.). Bei der Inzidenz liegt das Mammakarzinom mit 1,15 Mio.

Neuerkrankungen jährlich auf Platz 2 hinter dem Bronchialkarzinom (1,35 Mio.), nach Parkin et al. (Parkin et al., 2005).

Bei 66% der Frauen lassen sich keine Risikofaktoren für die Entstehung eines Mammakarzinoms ausmachen (Harris et al. 1992). Dennoch existiert eine Vielzahl an mehr oder weniger anerkannten Risikofaktoren. Hauptrisikofaktor für die häufigste Krebserkrankung der Frau ist das Lebensalter (Pike 1987). Bei der Brustkrebsentstehung geht man des Weiteren in 5 bis 10% der Fälle von einer familiären Disposition aus, die u. a. durch eine Mutation der Gene BRCA-1 und BRCA-2 (breast cancer susceptibility gene) hervorgerufen wird. Mutationen der Gene BRCA-1 und BRCA-2 werden zu 80 bis 90% für familiär bedingten Brustkrebs verantwortlich gemacht, weitere Brustkrebs-assoziierte Syndrome sind das Cowden-, Li-Fraumeni- und Peutz-Jeghers-Syndrom sowie die Ataxia teleangiectatica (Thull und Vogel 2004).

Diversen reproduktiven Faktoren wird eine Beeinflussung des Brustkrebsrisikos nachgesagt. Eine frühe Menarche, eine späte Menopause, Nulliparität und höheres Alter bei Erstgeburt werden demzufolge als Risikofaktoren angesehen, durch Multiparität und längeres Stillen kann das Karzinomrisiko reduziert werden (Kelsey et al. 1993). Eine insgesamt gesunde

Lebensführung mit ausreichender körperlicher Bewegung, Vermeidung von Übergewicht und übermäßigem Alkoholgenuss sollen sich ebenfalls begünstigend auswirken (Carmichael et al. 2009).

Kontrovers diskutiert wird die Rolle der Hormonersatztherapie (HRT) als Risikofaktor. Eine groß angelegte britische Studie wies eine höhere Tumorzinzidenz durch die HRT, insbesondere bei Östrogen-Progesteron-Kombinationen, nach (Beral 2003, Gray 2003). Auch die orale Kontrazeption scheint das Brustkrebsrisiko etwas zu erhöhen. Nach zehnjähriger Einnahmepause soll sich das Risiko jedoch wieder auf dem Niveau der Frauen ohne Kontrazeption bewegen (Collaborative Group 1996, Deligeoroglou et al. 2003).

Diverse architektonische Störungen des Brustgewebes neigen zur Entartung, darunter bestimmte Formen mastopathischer Veränderungen mit Zellatypien, Papillomatosen mit Zellatypien und atypische duktale sowie lobuläre Hyperplasien (Bässler 1997).

Neben den bereits Genannten tauchen in der Fachliteratur ständig neue vermutete Risikofaktoren auf. Aktuell wird beispielsweise die Melatonin-Hypothese diskutiert, derzufolge Melatonin einen protektiven Einfluss bei Krebsarten wie dem Mamma- und Prostata-Karzinom haben soll. Durch verkürztes Schlafverhalten oder Schichtarbeit wird die Melatoninproduktion unterdrückt, die jedoch auch im Alter nachlässt. Das arbeitsmedizinische Institut der Universität Köln fand in einer Metaanalyse von 30 Studien ein um 60% erhöhtes Brustkrebsrisiko bei Frauen, welche die angesprochenen Schlafgewohnheiten aufwiesen (Erren et al. 2009).

2.2 Veränderungen der Mamma

2.2.1 Bösartige Veränderungen

Häufigste Ausprägung einer bösartigen Veränderung der Mamma ist das invasive duktale Karzinom (IDC), das vom Gangepithel der Brustdrüse ausgeht. Weniger häufig ist das von den Azini der Drüsenläppchen ausgehende invasive lobuläre Karzinom (ILC). Wesentlich seltener und von meist besserer Prognose sind tubuläres, muzinöses, papilläres sowie medulläres Karzinom. Mischformen insbesondere mit invasiv-duktalem Karzinomen können ebenfalls auftreten.

Abzugrenzen sind nichtinvasive Karzinomvorläufer wie das duktale Carcinoma in situ (DCIS) und das lobuläre Carcinoma in situ (LCIS), auch hier befindet sich die Lokalisation jeweils in den Duktuli bzw. Drüsenläppchen.

Eine besondere Erscheinungsform stellt das inflammatorische Mammakarzinom dar. Neben klassischen Entzündungszeichen wie Rötung und Hyperthermie ist ein Befall der Hautlymphgefäße charakteristisch. Eine Abgrenzung zur Mastitis non-puerperalis gelingt teils nur mittels histologischer Abklärung.

Weiterhin existieren zahlreiche äußerst seltene Karzinomformen, vereinzelt treten auch primäre Lymphome der Brust oder Weichteilsarkome wie beispielsweise das Cystosarcoma phylloides auf. In der Mamma können auch metastatische Absiedlungen von Tumoren anderer Organe auftreten (Bässler 1997).

2.2.2 Gutartige Veränderungen

Differentialdiagnostisch sind unterschiedliche gutartige Neoplasien vom Mammakarzinom abzugrenzen. Am häufigsten handelt es sich dabei um das Fibroadenom, das Papillom, die fibrozystische Mastopathie mit oder ohne Zellatypien und das Cystadenoma phylloides. Papillomatose mit atypischen Epithelproliferationen und fibrozystische Mastopathie mit ausgedehnten Zellatypien gehen mit einer erhöhten Entartungswahrscheinlichkeit einher und werden der Gruppe der gutartigen Läsionen mit unklarem biologischen Potenzial zugerechnet. Die Grenzen zwischen atypischer duktaler Hyperplasie (ADH) und DCIS sind fließend (Bässler 1997).

2.3 Mammadiagnostik

2.3.1 Klinische Untersuchung

Neben ärztlichem Anamnesege­spräch stehen Inspektion und Palpation beider Mammae sowie die Untersuchung der axillären, infra- und supraklavikulären Lymphabflussgebiete am Anfang der klinischen Diagnostik (Kreienberg 2003). Neben der ärztlichen Untersuchung existiert auch die Empfehlung, alle Frauen in der monatlichen Selbstuntersuchung zu schulen (Albert und Schulz 2003), da die Mehrheit der Karzinome durch die Patientinnen selbst entdeckt wird. Aufgrund der vermehrten Anstrengungen durch das Mammographie-Screening scheint der Anteil an primär selbst erkannten Karzinomen jedoch zu sinken. Wurden ursprünglich 72% der Mammakarzinome durch die betroffenen Frauen selbst entdeckt (Schleicher und Ammon 1998), so berichtet eine aktuellere Studie über einen Anteil von lediglich 48,9% (Guth et al. 2008). Selbige Arbeit zeigt auch die Grenzen der klinischen Untersuchung auf. So lag der Median der getasteten Tumorgröße bei Ersterkennung über den Verlauf der Zeit trotz verstärkter gesundheitspolitischer Kampagnen und gesteigerten öffentlichen Interesses unverändert bei 21 mm, während die radiologisch erkannten Tumoren eine mediane Größe von 12 mm aufwiesen. Einige Studien gehen sogar so weit, dass sie den Wert der Selbstpalpation vollständig widerlegen. So konnte durch die regelmäßige Selbstuntersuchung keine Mortalitätsreduktion bewiesen werden, vielmehr häuften sich Biopsien gutartiger Befunde (Thomas et al. 2002).

2.3.2 Mammographie

Die Mammographie ist das wichtigste bildgebende Verfahren in der Mammadiagnostik, bei dem die Brustdrüsen und ihre axillären Ausläufer untersucht werden. Bei noch intaktem Menstruationszyklus soll die Untersuchung in der ersten Zyklushälfte durchgeführt werden (Bock et al. 2003). Prämenstruell ist die Brust meist druckempfindlicher und gespannt, zudem könnten die hormonell bedingten Wassereinlagerungen falsche Knoten vortäuschen. Die gleiche Empfehlung bezüglich des Zeitpunktes der Durchführung gilt auch für die MRM (Müller-Schimpfle et al. 1997).

Die Sensitivität der Mammographie ist von mehreren Faktoren abhängig. Mit zunehmender Involution der Brust steigt die Krebsentdeckungsrate, so dass bei Frauen über 65 eine Sensitivität bis 94% gefunden werden konnte, während Frauen unter 40 eine Sensitivität von

54% aufwiesen (Kerlikowske et al. 1996, Rosenberg et al. 1998). Mit zunehmendem Alter der Patientin und Fortschreiten der Involution des Drüsenkörpers gewinnt die Mammographie somit an Aussagekraft (Saarenmaa et al. 2001). Als Vorteil gegenüber allen anderen angewandten diagnostischen Verfahren erweist sich die Detektion von Mikrokalk, dessen Struktur und räumliche Anordnung mammographisch beurteilt werden kann. Das Vorkommen suspekter Mikrokalzifikationen ist bei nicht-invasiven Karzinomformen häufiger, daher kommt der Erkennung von mammographischem Mikrokalk eine wichtige Rolle beim Aufspüren von Karzinom-Frühformen zu (Hermann et al. 1988).

Erschwerend für die Suche nach malignen Auffälligkeiten sind postoperative narbige Veränderungen sowie Brustimplantate bei Verlaufskontrollen zum Rezidivausschluss (Bone et al. 1995). Eine hormonelle Beeinflussung in Form einer HRT wirkt sich ebenfalls negativ auf die Diagnostik aus (Laya et al. 1996).

Trotz zahlreicher existierender Möglichkeiten der Mamma-Bildgebung ist die konventionelle Mammographie nach wie vor die bislang einzig anerkannte Screening-Methode für Frauen zwischen dem 50. und 70. Lebensjahr, wobei die Untersuchung in zweijährigen Intervallen erfolgt (Schreer und Engel 2003). Zahlreiche Studien wiesen einen Nutzen hinsichtlich der Mortalitätsreduktion infolge des Mammographie-Screenings nach. 1993 wurde eine Metaanalyse zu schwedischen, randomisierten Screening-Studien mit Röntgenmammographie publiziert, die bei 282.777 Frauen eine Mortalitätssenkung an Brustkrebs nachwies. Am wirksamsten erschien das Mammographiescreening in der Gruppe der Frauen zwischen 50 und 69 Jahren, wo man eine Senkung der Mortalität um 29% fand (Nystrom et al. 1993).

2.3.3 Sonographie

Die Mammasonographie ist ein wichtiges additives Verfahren neben klinischer Untersuchung und Mammographie. Domäne der Ultraschalltechnik ist dabei die Differenzierung zwischen solidem Befund und Zyste sowie die Unterstützung bei Anwendung interventioneller Methoden in Form von Punktionen oder operativen Drahtmarkierungen. Ebenfalls nützlich ist das Verfahren bei Abklärung unklarer Tastbefunde insbesondere bei mammographisch dichten Brüsten (Jackson 1995). Insgesamt wird eine sonographische Abklärung sämtlicher palpatorisch und mammographisch erhobener Herdbefunde empfohlen. Bei Mammographien mit Zuordnung der BI-RADS-Kategorien 4 oder 5, also solchen, bei denen nicht mit größter Sicherheit ein Malignom ausgeschlossen werden kann, ist die Mammasonographie obligatorisch. Durch die fehlende Strahlenbelastung ergibt sich auch ein großer Vorteil für

den Einsatz bei jungen Frauen, bei Schwangeren und während der Laktation (Madjar et al. 2003). Nachteilig ist ein starker Abfall der Sensitivität mit abnehmender Herdgröße, so dass die Ultraschalluntersuchung als Screening-Methode ungeeignet erscheint (Teh und Wilson 1998).

2.3.4 MR-Mammographie

Überblick über Geschichte und Prinzip der MRM

1946 wurde erstmals das Zeichen der „nuclear magnetic resonance“ beschrieben (Purcell et al. 1946, Bloch et al. 1946). Sie beruht auf den Eigenschaften von Atomkernen ungerader Nukleonenzahl, die über einen als Kernspin bezeichneten Eigendrehimpuls verfügen, der ein magnetisches Moment entstehen lässt. Von medizinischer Relevanz ist vor allem das Wasserstoffatom.

Der moderne MR-Tomograph erzeugt mit einem supraleitenden Magneten ein starkes Magnetfeld, aus dem ein Summationsmoment resultiert, das in die Hauptfeldrichtung zeigt (z-Achse). Zur Auslenkung des Summationsmomentes aus der Hauptachse (z-Achse) kommt es durch ein eingestrahltes magnetisches Wechselfeld über eine als Resonanz bezeichnete Energieübertragung. Nach Beendigung des Auslenkimpulses kommt es zur Rückführung in den Ursprungszustand, der zur Energieabgabe führt und gemessen werden kann. Die Zeit, innerhalb derer die Wasserstoffkerne in den Gleichgewichtszustand zurückkehren, heißt Relaxationszeit, von der bei der MRT klassischerweise zwei Zeiten näher bezeichnet werden, nämlich die T1- oder longitudinale sowie die T2- oder transversale Relaxationszeit. Die bei Rückkehr in den Ausgangszustand ausgestrahlte Energie wird durch Spulen empfangen, in denen folglich Ströme induziert werden. Ein Computer wertet die empfangenen Signale zur Bildrekonstruktion aus, wobei die Ortskodierung des empfangenen Signals auf künstlich mittels Gradientenspulen erzeugter Feldstärkegradienten beruht (Reiser et al. 2006). Erkenntnisse über unterschiedliche Relaxationszeiten von gesundem und Tumorgewebe wiesen 1973 erstmalig auf den potentiellen Nutzen der MRT in der onkologischen Diagnostik hin (Damadian et al. 1973). Mansfield et al. konnten dies durch Verifizierung anhand von Untersuchungen an frisch entnommenem Operationsmaterial auch auf die Mammadiagnostik übertragen (Mansfield et al. 1979). Nachdem erste Untersuchungen wegen schlechter Signal-Rausch-Verhältnisse zunächst wenig überzeugende Ergebnisse lieferten (Ross et al. 1982, El Yousef et al. 1984), gelang es in der Folgezeit, die Untersuchungen vermehrt an eine

zweckmäßige Bildgebung der Mamma anzupassen. Dazu gehörte erst die Entwicklung spezieller Mamma-Einzelspulen (Stelling et al. 1985), später setzte sich die Mamma-Doppelspule durch, so dass eine simultane Darstellung beider Mammae möglich wurde (Kaiser und Kess 1989). Weitere Vorteile der Doppelspule ergaben sich durch Zeitersparnis bei gleichzeitiger Aufnahme beider Mammae, atmungs- und herzschlagbedingte Artefakte der vorher verwendeten Ganzkörperspulen wurden minimiert. Auch die sich durchsetzende Untersuchung in Bauchlage ergab qualitativ bessere Ergebnisse, weil so die schwerkraftbedingte Gewebekompression vermieden wurde (Wolfman et al. 1986). Ein weiterer Meilenstein war die Entwicklung des MR-Kontrastmittels Gadolinium-Diethyltriaminpentaacetat (Gd-DTPA) (Carr et al. 1984). Mit dessen Hilfe konnte man sich tumorbiologische Eigenschaften in der Bilddarstellung zunutze machen. Man geht davon aus, dass die Sauerstoff- und Nährstoffversorgung von Tumoren aufgrund deren expansiven Wachstums ab einem gewissen Grad nicht mehr durch Diffusion allein gewährleistet werden kann. Der resultierende hypoxische Stress führt zur Ausschüttung von Wachstumsfaktoren wie *vascular endothelial growth factor* (VEGF), was zur Aussprossung bestehender und Bildung neuer Kapillaren führt. Dabei kommt es auch zu einer grundsätzlich veränderten Gefäßarchitektur mit kapillaren Lecks und arteriovenösen Shunts, welche in vermindertem Maße der physiologischen Perfusionsregulation unterliegt (Folkman und Klagsbrun 1987, Folkman et al. 1989, Folkman 1992, Schneider und Miller 2005). Dieser als Angiogenese oder Neoangiogenese bezeichnete Prozess bildet die Grundlage für die Darstellung auch kleinster Neoplasien. Es wird von einer potenziell möglichen Detektion von Tumoren ab 3 mm Größe berichtet (Kaiser und Zeitler 1989).

Die Anwendung schneller Gradientenecho-Sequenzen ermöglichte zudem die repetitive Darstellung der gleichen Schichten vor und nach KM-Applikation (Haase et al. 1986). Da anreichernde Läsionen und umgebendes Fettgewebe ähnliche Signalintensitäten aufweisen können, wurden durch Subtraktionstechniken nicht anreichernde Strukturen eliminiert (Kaiser und Zeitler 1984, Haase et al. 1986). Diese neu entwickelten technischen Modalitäten konnten erfolgreich auf die MR-mammographische Bildgebung übertragen werden (Kaiser und Zeitler 1989). In den USA wurden dagegen Fettsättigungssequenzen entwickelt (Harms et al. 1993b).

Zahlreiche Techniken wurden und werden in der MR-Mammographie verwendet. Die diesbezüglichen Publikationen weisen zum Teil frappante Mängel an methodischen Angaben auf, allerdings konnten in einer aktuellen Metaanalyse keine prinzipiellen diagnostischen

Unterschiede zwischen verschiedenen Untersuchungstechniken identifiziert werden (Warren et al. 2009).

Eine Vielzahl an Publikationen hat belegt, dass die MR-Mammographie ein hochsensitives Verfahren zur Detektion von Brustkrebs ist. So werden Sensitivitäten für einen Bereich von 82,5 bis 100% ausgewiesen (Kaiser 1993, Fobben et al. 1995, Stomper et al. 1995, Heiberg et al. 1996). Die Spezifität wird mit Werten zwischen 37% und 97% angegeben (Harms et al. 1993a, Kaiser 1993). Heywang-Köbrunner und Altland betonen in der S-3-Leitlinie *Brustkrebs-Früherkennung in Deutschland* den inversen Zusammenhang von Sensitivität und Spezifität, welche von der Definition der Malignitäts- und Benignitätskriterien abhängig sind (Heywang-Köbrunner und Altland 2003). Eine große Metaanalyse beziffert die Sensitivität mit 90% und die Spezifität mit 72%. Diese Kennwerte waren, vergleichbar zu der Feststellung in der angesprochenen S-3-Leitlinie, je nach untersuchter Studie abhängig von der Prävalenz bösartiger Neubildungen sowie davon, wie viele und welche Kriterien zur Malignitätsdeutung herangezogen wurden (Peters et al. 2008). Nachteile der MRM sind die höhere Kostenintensivität und die fehlende Möglichkeit der Mikrokalkdetektion.

Aktuell ist die MRM speziellen Fragestellungen vorbehalten. Die European Society of Breast Imaging (EUSOBI) empfiehlt folgende Indikationen für die Durchführung einer MRM (Mann et al. 2008):

- Screening-Untersuchung für Hochrisikopatientinnen wie Mutationsträgerinnen der Gene BRCA-1 und BRCA-2
- Abklärung von in der konventionellen Bildgebung gefundenen unklaren Vorbefunden
- Screening der kontralateralen Brust bei histologisch nachgewiesenem unilateralem Brustkrebs, aber auch präoperatives Staging
- Bildgebung der Brust bei Metastasen eines unbekannten Primärtumors (CUP-Syndrom)
- Monitoring bei neoadjuvanter Chemotherapie
- Rezidivausschluss bei Z. n. brusterhaltender Therapie
- MR-basierte Biopsie

2.3.5 Invasive Verfahren zur histologischen Diagnostik

Verfahren zur Gewinnung von Zellmaterial oder Gewebeproben sind u. a. die Feinnadel-aspirationszytologie (FNB), die Stanzbiopsie und die Vakuumbiopsie. Sie können sowohl unter sonographischer Sicht als auch stereotaktisch gesteuert erfolgen. Voraussetzung für eine sonographisch gesteuerte Biopsie ist die sichere Identifikation mit Ultraschall. Ungeeignet ist die Sonographie für die Abklärung mammographisch sichtbaren Mikrokalks. Eine Studie konnte beispielsweise nur in 9% der Fälle bei mammographisch dargestelltem Mikrokalk einen Befund im Ultraschall zuordnen (Rissanen et al. 1998). Bei der stereotaktisch geführten Gewebeentnahme ist eine Aufnahme in zwei Ebenen notwendig, um dreidimensionale Koordinaten für den Sitz des Tumors zu erhalten. Hauptindikation für die Stereotaxie ist die Abklärung mammographisch sichtbaren Mikrokalks. Mittels Präparateradiographie lassen sich schließlich Verkalkungen bei korrekter Entnahme dokumentieren.

Ogleich Stanz- und Vakuumbiopsie anerkannte Verfahren hoher diagnostischer Genauigkeit zur Abklärung suspekter Raumforderungen sind, wurden die Unterschätzungsraten für bestimmte histologische Subgruppen durch zahlreiche Publikationen belegt. Diese sagen aus, dass bei histologischer Diagnose einer ADH oder eines DCIS nach Stanzbiopsie bei histopathologischer Aufarbeitung nach erfolgter OP zu einem Prozentsatz von ca. 30% auch invasive Tumorzellen auffindbar waren (Schueller et al. 2008, Youk et al. 2008).

Die FNB hat sich als Standardmethode u. a. wegen teils problematischer Materialgewinnung nicht durchsetzen können. Etabliert hat es sich hingegen, auffällige Befunde stanzbiopsisch abzuklären, wobei Sensitivitäten von nahezu 100% erreicht werden (Schulz-Wendtland et al. 1997, Crystal et al. 2005). Mit der Vakuumbiopsie lassen sich etwas bessere Ergebnisse als mit der Stanzbiopsie erreichen (Fahrbach et al. 2006).

2.4 Beschreibung MR-mammographisch sichtbarer Läsionen

Im Zuge der Entwicklung der MRM-Technik wurden diverse Deskriptoren für MR-mammographische Anreicherungen entwickelt, welche sich darin unterscheiden, dass sie die Läsion entweder morphologisch charakterisieren oder auf die Kontrastmittelkinetik eingehen. Hinsichtlich der Beschreibung des KM-Anreicherungsverhaltens wurde zunächst die Geschwindigkeit der KM-Aufnahme bewertet, konsekutiv wurde eine rasche KM-Aufnahme als wichtiges Malignitätsmerkmal definiert (Kaiser 1989). Weiterhin erfolgte die

Beschreibung des Kontrastmittelanreicherungsmusters innerhalb des Herdes. Vor allem eine initiale Anreicherung der Herdperipherie mit im zeitlichen Verlauf zunehmender Anreicherung in Richtung der Herdmitte stellt sich als hochprädiktiv für ein Mammakarzinom dar (Buadu et al. 1997, Liberman et al. 2002). Aktuell kaum wegzudenken ist die Beobachtung der zeitlichen Kontinuität der Signalintensität des fraglichen Befundes. Als hochsuspekt gilt dabei eine initial starke Anreicherung innerhalb der ersten zwei Minuten mit nachfolgendem Signalabfall, während ein schwächerer und im weiteren Verlauf anreichernder Befund eher für einen benignen Prozess spricht (Kaiser 1989, Kaiser 1992, Kuhl et al. 1999b). Aber auch zahlreicher morphologischer Deskriptoren bedient man sich bei der Herdbefundung. Im Blickpunkt steht hierbei die Beurteilung von Form und Berandung. Runde, ovale oder lobulierte Läsionen sowie eine scharfe Abgrenzung zum umgebenden Parenchym sprechen für einen benignen Prozess, irreguläre Form im Sinne einer irregulären, sternförmigen oder retikulären Ausprägung und eine schlechte Abgrenzbarkeit oder radiäre Ausziehungen des Herdes, in der Literatur als „spikuliert“ bezeichnet, deuten auf ein Malignom hin (Nunes et al. 1997b, Fischer et al. 1999, Liberman et al. 2002). Als weiteres morphologisches Kriterium wurde das für Malignome untypische Vorhandensein von Septen evaluiert (Nunes et al. 2001, Erguvan-Dogan et al. 2006).

Aufgrund der Uneinheitlichkeit der MRM-Befundung gab es immer wieder verschiedene Bestrebungen, im Sinne einer Qualitätssicherungsmaßnahme ein Lexikon für die Beschreibung von Läsionen zu entwickeln. Am Vorbild eines vergleichbaren Vorgehens bei der Röntgenmammographie wurde dies durch das ACR im Jahre 2003 erstmals publiziert, wobei ein Großteil oben genannter Deskriptoren einbezogen und definiert wurden (ACR 2003).

Zu den bereits Erwähnten wurden weitere MR-Zeichen untersucht. Als zusätzliches dynamisches und malignomtypisches Kriterium wurde das Blooming definiert, eine im zeitlichen Verlauf zunehmende Unschärfe der Berandung eines Herdes bei initial scharfer Abgrenzbarkeit (Fischer et al. 2004).

Als morphologische Parameter wurden weiterhin das sog. Root- und das sog. Hook-Zeichen eingeführt. Beide dienen der Beschreibung von Herdausläufern, ersteres meint dabei jeglichen Ausläufer, das Hook-Zeichen ist definiert durch jegliche Verbindung eines Herdes zum M. Pectoralis. Beide Zeichen sind stark prädiktiv für Malignität (Malich et al. 2005, Kaiser 2007).

Des Weiteren wurden auch Zeichen beschrieben, welche das die Läsion umgebende Gewebe charakterisieren, darunter das Vorhandensein von Ödemen (Malich et al. 2005), eine ipsilateral prominente Gefäßzeichnung (Sardanelli et al. 2005) bzw. ein zu dem Tumor

führendes Gefäß (Malich et al. 2005, Fischer et al. 2006), Cutisverdickung und Mamillensaumunterbrechung (Malich et al. 2005). Diese Zeichen gelten ebenfalls als Hinweis auf ein Malignom.

In Anlehnung an eine abschließende Beurteilung einer Mammographie wird auch für die MRM ein solches Vorgehen empfohlen. Abhängig von der vergebenen BI-RADS-Kategorie sollte ein in Tab. 2.1 wiedergegebenes Procedere erfolgen.

Tabelle 2.1: BI-RADS-Einstufung des ACR

BI-RADS-Kategorie	Erklärung	Karzinomrisiko	Procedere
0	Beurteilung unvollständig und nicht möglich	-	Weitere Bildgebung zur Gesamtbeurteilung notwendig
1	Kein Befund	0%	Fortführung Screening gemäß Leitlinien
2	Gutartiger Befund	0%	Fortführung Screening gemäß Leitlinien
3	Wahrscheinlich gutartiger, kontrollbedürftiger Befund	$\leq 2\%$	Verlaufskontrolle nach 6 Monaten
4	Suspekter, abklärungsbedürftiger Befund	2-90%	Sofortige histologische Abklärung
5	Hochgradiger Malignomverdacht	$> 90\%$	Sofortige histologische Abklärung
6	Histologisch gesichertes Malignom		Therapie

2.5 MR-mammographische Interpretationsmodelle

Aufgrund der häufig an der MRM kritisierten niedrigen Spezifität schufen zahlreiche Autoren mittels Kombination morphologischer und kinetischer Parameter diverse Interpretationsmodelle mit dem Ziel, die Spezifität zu erhöhen und besser zwischen malignen und benignen Läsionen unterscheiden zu können. Nachfolgend sei eine Auswahl der Vielzahl inzwischen publizierter Interpretationsmodelle gegeben.

Nunes et al. entwickelten 1997 erstmals einen Modellvorschlag, welcher vorwiegend morphologische Beurteilungskriterien einschloss und die Signalintensität des Herdes bewertete (Nunes et al. 1997a). U. Fischer et al. erstellten 1999 einen einfachen empirischen Score, der die maximale Signalintensität, den Signalintensitätskurvenverlauf, das KM-Anreicherungsverhalten, die Herdform und die Berandung berücksichtigte. Für jedes Merkmal wurde ein Punkt zwischen Null und zwei vergeben, anschließend die Summe errechnet und in Abhängigkeit von deren Höhe der Herd als gutartig oder bösartig beschrieben (Fischer et al. 1999). D. Fischer et al. erweiterten diesen Score mit von Malich und Mitarbeitern erarbeiteten Deskriptoren, die vor allem bei unsicheren Fällen die diagnostische Genauigkeit verbessern sollten (Malich et al. 2005, Fischer et al. 2005).

Methodisch ähnlich wie Nunes erstellte eine Arbeitsgruppe um Tozaki mittels Klassifikationsanalyse ein Baumdiagramm als einfache Entscheidungshilfe für die Dignitätsvorhersage, hier unter Nutzung von Informationen über Form, Berandung, T2-Signalintensität, Septen und Signalintensitätskurvenverlauf (Tozaki et al. 2005).

Schnall et al. entwickelten ein Interpretationsmodell mittels binärer logistischer Regression, das die Wahrscheinlichkeit für das Vorliegen von Malignität berechnen konnte. Im Unterschied zu den vorher genannten Modellen wurden hierbei aber auch Kriterien eingeschlossen, die nicht allein durch die MR-Bildgebung determiniert sind. Diese waren beispielsweise das Lebensalter oder das Vorhandensein von Mikrokalk in der Röntgenmammographie (Schnall et al. 2006).

Zusammenfassend ist bei der Vielzahl der Modelle festzustellen, dass ihre Vergleichbarkeit eingeschränkt ist. Dies ist zurückzuführen auf unterschiedliche Untersuchungsmethodik, abweichende Definitionen der Beurteilungskriterien, stark divergierende Fallzahlen und der teils fundamental verschiedenen Zusammensetzung der Untersuchungskollektive.

3 Ziele dieser Arbeit

Zahlreiche Vorarbeiten widmeten sich dem Ziel der systematischen Klassifikation von Herdbefunden anhand morphologischer und dynamischer Deskriptoren. Die verwendeten Ansätze waren dabei unterschiedlich und schlossen sowohl retrospektive als auch prospektive Analysen ein. In mehreren Arbeiten wurde die Übereinstimmung verschiedener Untersucher adressiert. Die verwendeten Ansätze reichten von empirischen über mathematische Methoden bis hin zu informatischen Methoden künstlichen Lernens. Ziel eines diagnostischen Modells ist die einheitliche Befunderstellung und Differentialdiagnostik unter Zuhilfenahme einheitlicher und klar definierter Deskriptoren. Voraussetzung ist eine genügend große Datenbank, aus der sich die Genauigkeiten der einzelnen deskriptiven Kriterien ermitteln lassen.

Methodisch ist durch Elimination redundanter, d. h. nicht abhängiger Deskriptoren eine Reduktion auf wenige Kriterien und so ein im klinischen Alltag akzeptables Modell zu erreichen.

Aufgrund obiger Überlegungen wurde in dieser Arbeit die Methode der binär-logistischen Regression gewählt. Diese ermöglicht die Parametrisierung einer multivariaten Verteilung, das heißt die konditionalen Wahrscheinlichkeiten einzelner Prädiktorvariablen lassen sich zu einer Gesamtwahrscheinlichkeit des Vorhandenseins eines binären Zustandes, hier dem Vorliegen einer malignen Neoplasie der Mamma, zusammenfassen. Mit Hilfe von Variablen-selektion soll ein möglichst einfaches Modell für die einheitliche Differentialdiagnostik von anreichernden Herdbefunden der Mamma unter Einschluss einer Vielzahl morphologischer und dynamischer Prädiktorvariablen unternommen werden. Das sich als Gesamtwahrscheinlichkeit des Vorliegens von Brustkrebs charakterisierte Ergebnis lässt zudem eine Abschätzung der diagnostischen Genauigkeit des ermittelten Modells sowie des Anteils und der Charakteristik der mit hoher Sicherheit klassifizierten Einzelbefunde zu.

Neben bereits etablierten Deskriptoren wie *Form*, *Rand*, *Septen*, *initiale KM-Anreicherung*, *postinitiale KM-Dynamik* und *internes KM-Anreicherungsverhalten* sollten auch neue Parameter in die Modellerstellung integriert werden.

Folgende konkrete Fragestellungen wurden bearbeitet:

1. Wie hoch ist die Prävalenz der einzelnen morphologischen und kinetischen MR-mammographischen Deskriptoren in benignen und malignen Herdbefunden?
2. Wie hoch ist die maximale mittels binär-logistischer Regression ermittelte Genauigkeit der zu analysierenden morphologischen und dynamischen Deskriptoren?
3. Lässt sich anhand einer geringen Anzahl von Deskriptorvariablen ein einfaches Regressionsmodell mit ausreichender Genauigkeit bezüglich der Differenzierung zwischen malignen und benignen Herdläsionen identifizieren?
4. Lässt sich ein Anteil von Herdbefunden anhand binärer logistischer Regression mit hoher Sicherheit klassifizieren, und wenn ja, bei welchen Konstellationen von Deskriptoren wäre dies der Fall?

4 Methoden

4.1 Einschluss-/ Ausschlusskriterien

Grundlage für die Datenerhebung bilden 4847 MR-Mammographien vom 01.12.1994 bis 31.03.2004 sowie 1109 derartige Untersuchungen aus der Zeit vom 03.01.2005 bis 16.10.2006, welche am Institut für Diagnostische und Interventionelle Radiologie (IDIR) des Universitätsklinikums Jena (UKJ) durchgeführt wurden.

Eine der Indikationen für die Durchführung der MRM waren die Frage nach Multifokalität, Multizentrität, Tumorausdehnung, Feststellung der Infiltration der Pectoralmuskulatur und die Untersuchung der kontralateralen Brust im Rahmen des präoperativen Stagings. Im Rahmen des Therapie-Monitorings erfolgte die Verlaufskontrolle unter neoadjuvanter Chemotherapie. Ein Einsatz erfolgte auch als Vorsorgeuntersuchung bei vorhandener genetischer Prädisposition und positiver Familienanamnese. War die Mamma unter Anwendung anderer konventioneller bildgebender Untersuchungsmethoden unzureichend beurteilbar, so wurde die Einschätzung der Dignität unklarer Herde mittels MRM versucht. Weitere Indikationen ergaben sich postoperativ für die Implantatkontrolle auf Lokalisation und Unversehrtheit sowie im Sinne einer Rezidivdiagnostik.

Kontraindikation für die MR-Untersuchung waren ferromagnetische Materialien im Körper wie Herzschrittmacher oder Cochlea-Implantate, KM-Allergie, Klaustrophobie und extreme Adipositas.

In die Studie eingeschlossen wurden ausschließlich Patientinnen, deren ambulante und/oder stationäre Betreuung durch die Klinik für Frauenheilkunde und Geburtshilfe des UKJ erfolgte. Patientinnen, bei denen in einem Zeitraum von weniger als 6 Monaten vor der Untersuchung eine Manipulation an der Mamma erfolgte, wurden ausgeschlossen. Dies war der Fall bei durchgeführter Bestrahlung oder Chemotherapie sowie jeglichem interventionellen Eingriff. Pathologisch anreichernde Läsionen fanden ebenfalls keine Berücksichtigung, wenn eine eindeutige histologische Zuordnung nicht möglich war.

Schließlich gelang aus den insgesamt 5956 MR-Mammographien die vollständige Erhebung der histopathologischen und der für diese Arbeit relevanten MR-mammographischen Daten bei 1012 Patientinnen im Alter zwischen 16 und 91 Jahren. Das Durchschnittsalter betrug

55,5 ± 13,1 Jahre, der Median ergab 55,0 Jahre. Bei den 1012 Patientinnen wurden 1084 anreichernde Prozesse histologisch verifiziert. Das Material zur feingeweblichen Aufarbeitung wurde bei 871 Läsionen durch eine Operation und bei den verbleibenden 213 per Stanzbiopsie gewonnen.

4.2 Einteilung der histopathologischen Befunde

Maßgeblich für diese Arbeit war, ob es sich bei einer in der MRM auffälligen Veränderung um einen gutartigen oder bösartigen Prozess handelte. Dementsprechend wurden sämtliche invasive oder nichtinvasive Karzinomformen, Lymphome, Weichteilsarkome oder Metastasen der Brust als maligne angesehen. Neoplasien wie Fibroadenome, Papillome, Cystadenomae phylloides, Entzündungen der Brustdrüse oder -gänge sowie jegliche Ausprägungen einer Mastopathie wurden als benigne betrachtet. Die Tab. 4.1 und 4.2 zeigen die Häufigkeiten der histopathologischen Subgruppen.

Nach obigen Kriterien wurden 1084 histologisch verifizierte Läsionen eingeschlossen. Die 648 malignen und 436 benignen Herdbefunde entsprachen dabei einer Häufigkeitsverteilung von 59,8% bösartigen und 40,2% gutartigen Befunden.

Die heterogene Gruppe der mastopathisch-proliferativen Veränderungen stellte mit 50,5% den Hauptanteil der gutartigen Befunde, gefolgt von den Fibroadenomen (23,6%) und Papillomen (19,0%). Vergleichsweise selten nahmen sich der Phylloidentumor (2,3%) und die entzündlichen Veränderungen der Mamma (4,6%) aus.

Tabelle 4.1: Häufigkeiten der histologischen Untergruppen gutartiger Befunde

Histologie	absolute Häufigkeit	relative Häufigkeit in %
Fibroadenom	103	23,6
Papillom	83	19,0
Cystadenoma phylloides	10	2,3
entzündliche Veränderungen	20	4,6
mastopathisch-proliferative Veränderungen	220	50,5

Bei den malignen Befunden dominierte mit einer Häufigkeit von 347 Fällen (53,5%) das invasive duktales Karzinom (IDC), gefolgt vom invasiven lobulären Karzinom (ILC), welches einen Anteil von 16,7% stellte. Ein Mischtyp aus IDC und ILC lag zu 3,7% vor. Ebenso unterschied man invasive Tumoren vom tubulären (4,2%), muzinösen (0,9%), medullären (0,5%) und papillären Typ (0,3%). Zu den sonstigen bösartigen Veränderungen fasste man andere seltene Karzinomformen (z.B. Plattenepithelkarzinom, neuroendokrines Karzinom), Metastasen jeglicher Art, Lymphome und maligne Weichgewebstumoren zusammen (insgesamt 7,3%).

Bei den nicht-invasiven Histologien fand man 69 DCIS (10,6%), 7 LCIS (1,1%) sowie 8 Mischtypen von DCIS und LCIS (1,2%).

Tabelle 4.2: Häufigkeiten der histologischen Untergruppen maligner Befunde

Histologie	absolute Häufigkeit	relative Häufigkeit in %
<i>invasive Formen</i>		
duktales Karzinom	347	53,5
lobuläres Karzinom	108	16,7
Karzinom mit duktalem und lobulären Komponenten	24	3,7
tubuläres Karzinom	27	4,2
medulläres Karzinom	3	0,5
muzinöses Karzinom	6	0,9
papilläres Karzinom	2	0,3
<i>Sonstige</i>	47	7,3
<i>nicht-invasive Formen</i>		
duktales Carcinoma in situ	69	10,6
lobuläres Carcinoma in situ	7	1,1
Carcinoma in situ mit duktalem und lobulären Komponenten	8	1,2

4.3 Untersuchungsablauf

Im Rahmen der Anamneseerhebung wurde die Patientin über zurückliegende Erkrankungen und Operationen vor allem gynäkologischer Art befragt. Des Weiteren wurden Informationen über Zahl und Verlauf von Schwangerschaften, onkologischen Erkrankungen bei nahen Verwandten, Medikamenteneinnahme mit Schwerpunkt auf Hormonersatzpräparate sowie Zyklusphase erfasst. Durch ein anschließendes Aufklärungsgespräch wurde die Patientin über den Untersuchungsablauf, deren Risiken und mögliche Kontrastmittelnebenwirkungen aufgeklärt. Um Unsicherheit und damit verbundene Bewegungsartefakte zu minimieren, wurde die Patientin auf die Lautstärke der Untersuchung hingewiesen und die Notwendigkeit dargelegt, während der gesamten Untersuchung still zu liegen. Im Anschluss erhielt die Patientin einen intravenösen Zugang über eine Kubitalvene oder, wenn dies nicht möglich war, über den Handrücken. Die Untersuchung wurde in Bauchlage durchgeführt und die Mammæ möglichst ohne Kompression und Deformation in die Doppelspule fallen gelassen. Die Untersuchungen erfolgten nach einem vom verwendeten Gerät abhängigen Messprotokoll (vgl. Tab. 4.4 bis 4.6).

Die Untersuchungen wurden an vier verschiedenen MR-Tomographen durchgeführt:

- Gyroscan S15 ACS II (Philips Medizintechnik, Hamburg, Deutschland)
- Philips Intera (Philips Medizintechnik, Hamburg, Deutschland)
- Siemens Vision (Siemens, Erlangen, Deutschland)
- Siemens Symphony (Siemens, Erlangen, Deutschland)

Tab. 4.3 demonstriert, wie viele Läsionen mit dem jeweiligen Gerät erfasst wurden. Im Zeitraum vom 1. Dezember 1994 bis 31. Dezember 2001 erfolgten die Untersuchungen dabei allesamt am Gyroscan S15 ACS II (Philips Medizintechnik, Hamburg, Deutschland). Danach kamen auch das Philips Intera (Philips Medizintechnik, Hamburg, Deutschland), das Siemens Vision (Siemens, Erlangen, Deutschland) oder das Siemens Symphony (Siemens, Erlangen, Deutschland) zum Einsatz. Im Zeitraum vom 01.10.2005 bis 16.10.2006 wurden die Aufnahmen ausschließlich vom Siemens Symphony (Siemens, Erlangen, Deutschland) angefertigt.

Tabelle 4.3: Anzahl der von den unterschiedlichen Tomographen erfassten Läsionen

MR-Tomograph	Erfasste Anreicherungen
Philips Gyroscan	649
Philips Intera	24
Siemens Vision	54
Siemens Symphony	357

Sämtliche Geräte waren mit einem supraleitenden Magneten der Feldstärke 1,5 Tesla (T) ausgestattet. Bei den jeweiligen MR-Tomographen wurde eine Mamma-Oberflächen-Doppelspule mit Empfangsfunktion der Firmen Philips (Philips Medizintechnik, Hamburg, Deutschland) bzw. Siemens (Siemens, Erlangen, Deutschland) verwendet.

Der prinzipielle Aufbau des Messprotokolls begann mit einer schnellen T1-gewichteten Localizer Sequenz. An dieser erfolgte die Planung T1-gewichteter Sequenzen transversaler Orientierung. Als nächstes kam eine T1-gewichtete native 2D-Gradientenechosequenz in koronarer Schnittführung zum Einsatz zwecks Beurteilung von Artefakten und Inhomogenitäten. Es schloss sich die transversale Kontrastmitteldynamik in einminütigen Abständen nach intravenöser Druckinjektion (3 ml/min) von 0,1 mmol/kg KG Gd-DTPA (Bayer Healthcare, Leverkusen, Deutschland) an. Bei Injektion in die Ellenbeuge waren der KM-Gabe 20 ml physiologische Kochsalzlösung zugesetzt, bei Injektion in die Handrückenvene waren es 30 ml. Die Schichtdicke variierte bei Ablauf der dynamischen Sequenz geräteabhängig zwischen 3 und 4 mm, die Schichtanzahl lag ebenfalls in Abhängigkeit vom verwendeten Tomographen zwischen 24 und 33. Nach der Kontrastmitteldynamik erfolgte eine koronare T1-gewichtete Sequenz, im Anschluss daran eine T2-gewichtete Turbo-Spin-Echo-Sequenz.

Dynamische Aufnahmen thoraxwandnaher Herde wurden alternativ in koronarer Schnittführung durchgeführt. Die Tabellen 4.4 bis 4.6 geben eine Zusammenfassung der gerätespezifischen Untersuchungsprotokolle.

Tabelle 4.4: Messprotokoll der MR-Mammographie mit den Philips-Geräten

	1.	2.	3.	4.	5.
Wichtung	T1	T1	T1	T1	T2
Sequenz	SE	FFE	FFE	FFE	TSE
Orientierung	transversal	koronar	transversal	koronar	transversal
Repetitionszeit	121	96	97	96	4000
TR in ms					
Echozeit TE	13	5	5	5	300
in ms					
Flipwinkel	90	80	80	80	90
in °					
Schichtdicke	5	4	4	4	4
in mm					
Distanzfaktor	10	10	10	10	10
in %					
Field of view	450	350	350	350	350
in mm					
Anzahl der Schichten	5	24	24	24	24
Basisauflösung	256	256	256	256	256
in Pixel					

Tabelle 4.5: Messprotokoll der MR-Mammographie mit dem Siemens Vision

	1.	2.	3.	4.	5.
Wichtung	T1	T1	T1	T1	T2
Sequenz	SE	FLASH	FLASH	FLASH	TSE
Orientierung	transversal	koronar	transversal	koronar	transversal
Repetitionszeit	220	220	220	220	12288
TR in ms					
Echozeit TE	4,6	4,66	4,66	4,66	259
in ms					
Flipwinkel	75	75	75	75	90
in °					
Schichtdicke	4	4	4	4	4
in mm					
Distanzfaktor	10	10	10	10	10
in %					
Field of view	350	350	350	350	350
in mm					
Anzahl der Schichten	24	24	24	24	24
Basisauflösung	256	256	256	256	256
in Pixel					

Tabelle 4.6: Messprotokoll der MR-Mammographie mit dem Siemens Symphony

	1.	2.	3.	4.	5.
Wichtung	T1	T1	T1	T1	T2
Sequenz	FLASH GRAPPA®	FLASH GRAPPA®	FLASH GRAPPA®	FLASH GRAPPA®	TSE
Orientierung	transversal	koronar	transversal	koronar	transversal
Repetitionszeit TR in ms	112	113	113	113	8900
Echozeit TE in ms	4,5	4,56	4,56	4,56	207
Flipwinkel in °	80	80	80	80	90
Schichtdicke in mm	3	3	3	3	3
Distanzfaktor in %	10	10	10	10	10
Field of view in mm	350	350	350	350	350
Anzahl der Schichten	33	33	33	33	33
Basisauflösung in Pixel	256*	256*	256*	256*	256*

*Die Pixelzahl konnte optional zugunsten einer konstanten Bildqualität bei jeder Untersuchung im MR-Tomographen Siemens Symphony individuell angepasst werden.

4.4 Auswertung und Datenerfassung

Die Beurteilung der MR-Aufnahmen erfolgte durch in der magnetresonanztomographischen Diagnostik der Mamma in der gleichen Institution einheitlich geschulte Radiologen. Die Untersuchungsreihenfolge unterlag dabei dem Zufall. Über Informationen bezüglich der histologischen Diagnose verfügten die befundenden Ärzte nicht. Anamnese und klinische Befunde von Voruntersuchungen waren hingegen wie in der klinischen Routine üblich bekannt. Die Beschreibung der kategorialen Deskriptorvariablen (auch als Deskriptoren, Kriterien oder Zeichen aufgeführt) einer anreichernden Struktur erfolgte im Konsensus. Vor dem Beginn der Studie wurden die zu analysierenden Kriterien gemeinsam und einheitlich definiert. Eine einwöchige Trainingsphase ging der Evaluation der Deskriptoren voraus. Für die Teilnahme an der Datenevaluation war eine Erfahrung in der Befundung von mindestens 500 MR-Mammographien obligatorisch. Aufgrund des langen Studienzeitraumes fand die Datenerfassung zu mehreren Zeitpunkten statt.

4.4.1 Beurteilung der morphologischen Kriterien

In Tab. 4.7 sind die Vergabemöglichkeiten morphologischer Läsionscharakteristika im Überblick dargestellt.

Bei der Anreicherungsform erfolgte eine Differenzierung in *rund/oval*, *lobuliert* oder *irregulär*. Die Berandung wurde als *scharf* oder *unscharf* bezeichnet.

Die Binnenstruktur des fraglichen Befundes im nativen T1-gewichteten Bild konnte als *homogen* oder *inhomogen* bewertet werden.

Schließlich konnte man bei der Beurteilung der Signalintensität nativer T1w-Bilder zwischen *hell*, *isointens* und *dunkel* wählen. Ebenso erfolgte eine Evaluierung der Helligkeit in der T2-Wichtung, wobei man zwischen *hyperintens* und *isointens* bzw. *hypointens* unterschied, jeweils in Bezug auf die Signaldarstellung zum Mammaparenchym.

Das Auftreten von Septen wurde zuerst in T1-gewichteten Nativ-Aufnahmen beurteilt. In den Subtraktionsaufnahmen wurde ein Anreichern der Septen überprüft, letztendlich wurden auch T2-Aufnahmen für die Beurteilung auf das Vorhandensein von Septen betrachtet. Schließlich wurde festgehalten, ob Septierungen auszumachen waren, und wenn ja, ob diese anreicherten oder nicht.

Jegliche Ausläufer eines Herdes wurden als positives *Root-Zeichen*, jegliche Verbindung zum M. pectoralis als positives *Hook-Zeichen* bewertet (Kaiser 2007) (siehe Abb. 4.1).

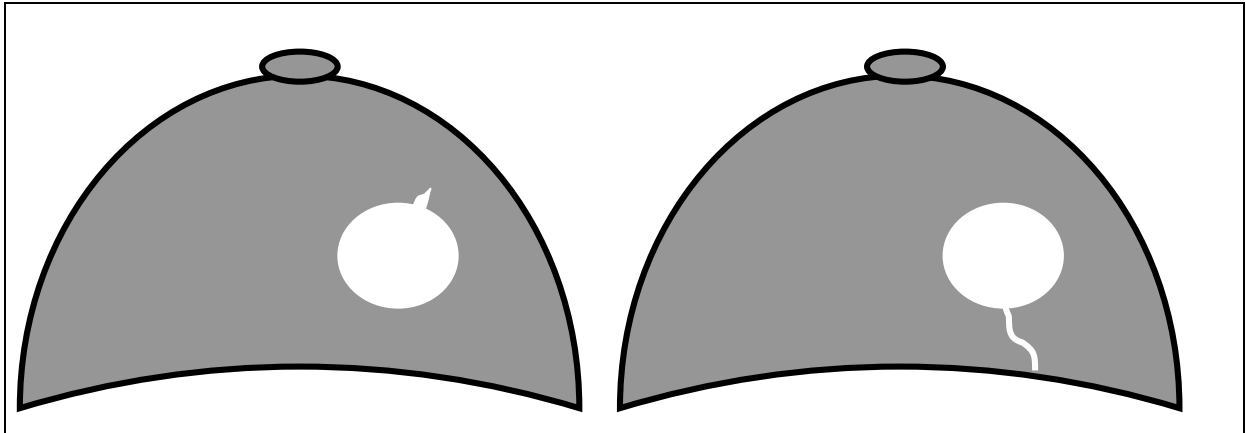


Abbildung 4.1: Root- und Hook-Zeichen

Schematische Darstellung von Root-sign (links) als beliebiger Herdausläufer und Hook-sign (rechts) als Ausläufer, der den M. pectoralis infiltriert

Tabelle 4.7: Deskriptoren der Läsionsmorphologie

Deskriptor	Ausprägung
Rand in T1-Subtraktion	scharf
	unscharf
Form in T1-Subtraktion	rund/oval
	lobuliert
	irregulär
Binnenstruktur in T1w nativ	homogen
	inhomogen
Signalintensität in T1w nativ	hell
	isointens
	dunkel
Septen	keine
	vorhanden
	anreichernd
Signalintensität in T2w	hell
	isointens/dunkel
Root-Zeichen	nein
	ja
Hook-Zeichen	nein
	ja

4.4.2 Beurteilung der dynamischen Kriterien

Nach der morphologischen Beurteilung erfolgte die Bewertung der dynamischen Kriterien der Läsion. Zunächst wurde das räumliche Kontrastmittelanreicherungsmuster (nachfolgend auch synonym als *internes Anreicherungsverhalten* oder *Binnenanreicherungsmuster* bezeichnet) beschrieben, welches *homogen*, *inhomogen*, *zentrifugal* oder *zentripetal* sein konnte. Beim *zentrifugalen Enhancement* lag der Schwerpunkt des Anreicherungsverhaltens zunächst im Zentrum, wobei im Zeitverlauf die Peripherie zunehmend anreicherte. Im Gegensatz dazu reicherte beim *zentripetalen Enhancement* initial die Peripherie an, und zentralere Bereiche kamen im zeitlichen Verlauf verstärkt zur Darstellung. Trafen diese MR-Zeichen nicht zu, war die Läsion entsprechend der Wahrnehmung in der Bildgebung homogen oder inhomogen anreichernd (Kaiser 2007) (Abb. 4.2).

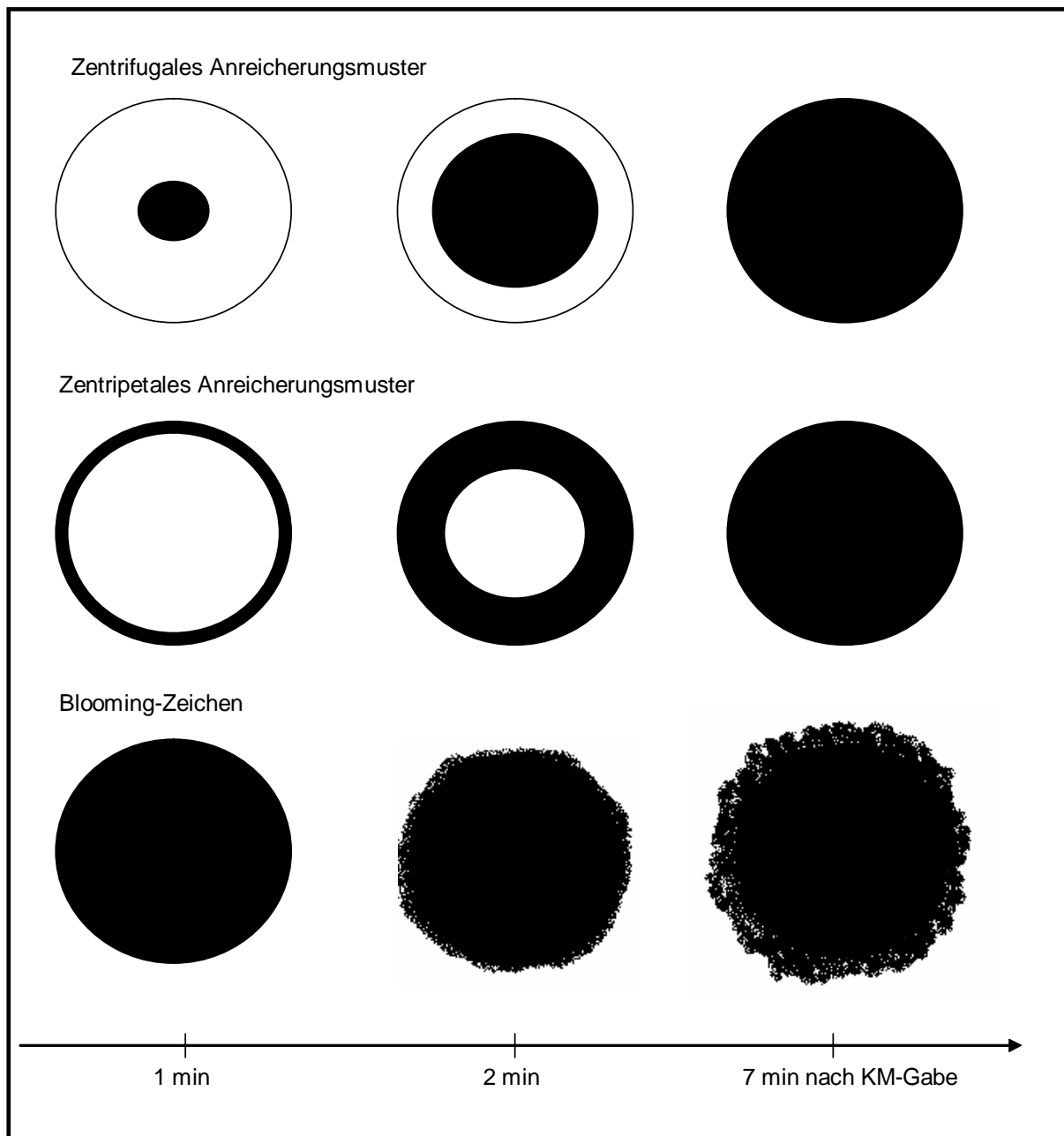


Abbildung 4.2: Schematische Darstellung KM-dynamischer Zeichen

Die zeitliche Konstanz des KM-Anreicherungsverhaltens wurde ebenfalls visuell festgehalten und für eine Dauer bis sieben Minuten nach KM-Applikation betrachtet (Abb. 4.3). Differenzierungsmöglichkeiten waren ein *kontinuierlicher (persistierender) Signalanstieg*, *Plateau* oder *wash-out*. Das *Wash-out-Phänomen* war definiert als Abschwächung der Signalintensität durch KM-Auswaschung nach initial schneller KM-Anreicherung (Kaiser 1992, Bick 2000). Das Sistieren der KM-Anreicherung nach Erreichen eines

Signalmaximums in den ersten Minuten wurde als *Plateau-Phänomen* bezeichnet (Kaiser und Zeitler 1989). Die unterschiedlichen Ausprägungen des Signalintensitätszeitverlaufes sind der Abb. 4.3 zu entnehmen.

Als weiteres dynamisches Kriterium wurde das Vorhandensein des sogenannten *Blooming-Zeichens* untersucht (Abb. 4.2). Es galt als positiv bei einer zunehmenden Unschärfe der Berandung eines anreichernden Herdes sieben Minuten nach KM-Injektion bei initial scharfer Berandung innerhalb der ersten Minute nach KM-Verabreichung sowie KM-Anreicherung während der ersten zwei Minuten nach Bolusinjektion (Fischer et al. 2004). Die Beurteilung der dynamischen Kriterien ist in Tab. 4.8 zusammengefasst.

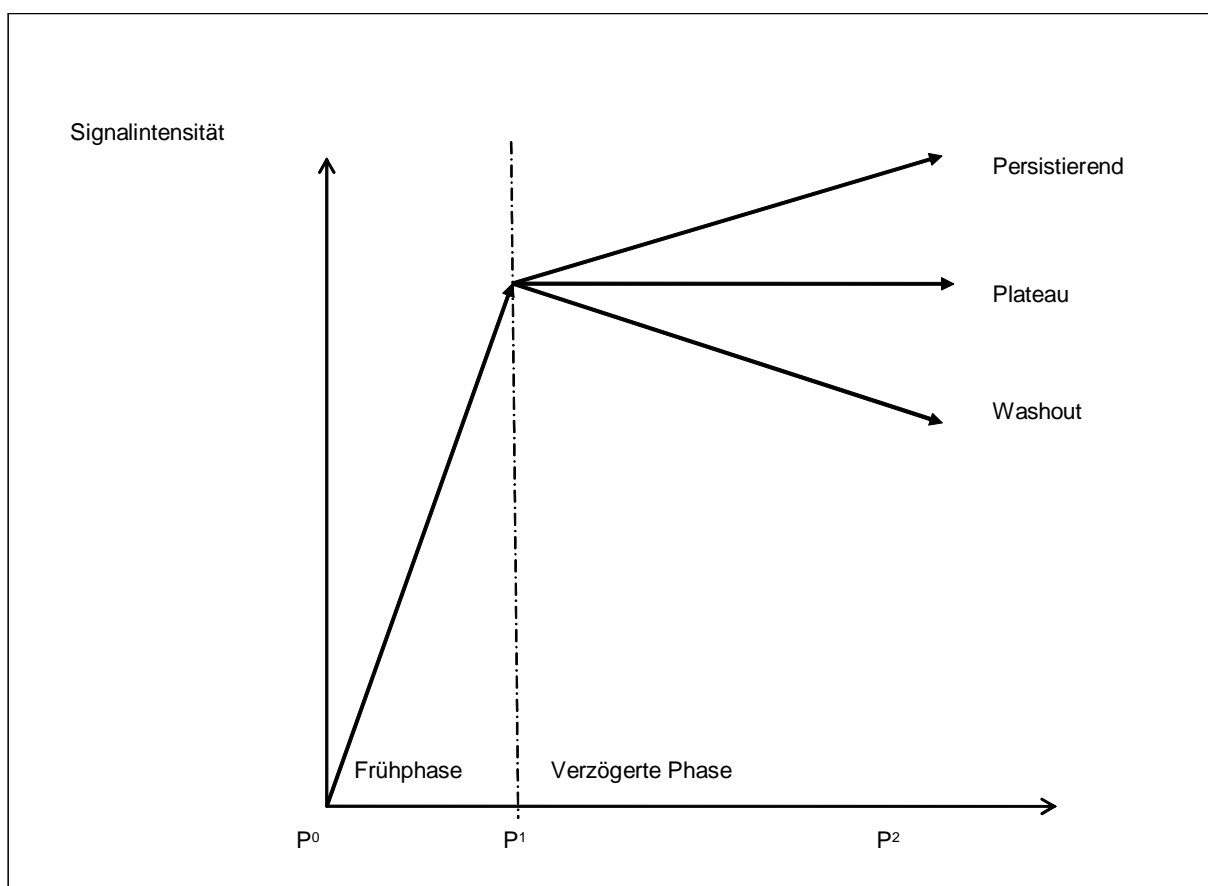


Abbildung 4.3: Schematische Darstellung des zeitlichen KM-Anreicherungsverhaltens

Tabelle 4.8: Beurteilungsmöglichkeiten für dynamische Deskriptoren

Deskriptor	Ausprägung
Räumliches Kontrastmittelanreicherungsverhalten	homogen inhomogen zentrifugal zentripetal
Visuelle Beurteilung der Konstanz der Signalintensität 7 min nach der KM-Injektion	kontinuierlicher Anstieg Plateau wash-out
Blooming unter Betrachtung der Subtraktionen bis 7 min nach der KM-Injektion	ja nein

4.4.3 Beurteilung des die Läsion umgebenden Gewebes

Neben der Beschreibung der morphologischen und KM-dynamischen Eigenschaften der Läsion wurden auch Strukturen und Merkmale berücksichtigt, welche die Läsion umgaben.

Im T2w-Bild erfolgte die Begutachtung auf Vorliegen eines Ödems im Mamma-Parenchym. Ein solches lag vor, wenn signalreiche Strukturen in T2-gewichteten Sequenzen weder Zysten, noch intraduktaler Flüssigkeit zugeordnet werden konnten.

Bei Lokalisation eines Ödems in unmittelbarer Umgebung des zu untersuchenden Befundes erfolgte eine Einteilung zugunsten eines *perifokalen Ödems* (Abb. 4.4). Beschränkten sich ödematöse Veränderungen nicht lediglich auf die Läsionsperipherie, so wurde die Eigenschaft *diffuses ipsilaterales Ödem* zugewiesen (Malich et al. 2005).

In Postkontrastaufnahmen wurde das Vorhandensein einer generellen ipsilateralen Hypervaskularisation beziehungsweise eines in die Läsion führenden Gefäßes geprüft (Sardanelli et al. 2005, Fischer et al. 2006) (siehe Abb. 4.4).

Eine ipsilaterale *Cutisverdickung* wurde ab einem Durchmesser von 3 mm als positiv bewertet. Zuletzt beobachtete man, ob eine Unterbrechung des nicht-anreichernden Mamillensaums durch anreichernde Abschnitte festzustellen war (Malich et al. 2005) (vgl. Tab. 4.9).

Tabelle 4.9: Deskriptoren der die Läsion umgebenden Auffälligkeiten

Deskriptor	Ausprägung
Ödem in T2	nein diffus ipsilateral perifokal
Gefäßauffälligkeiten in T2 oder Dynamik	nein ipsilateral prominent zuführend
Cutisverdickung	nein ja
Mamillensaumunterbrechung	nein ja

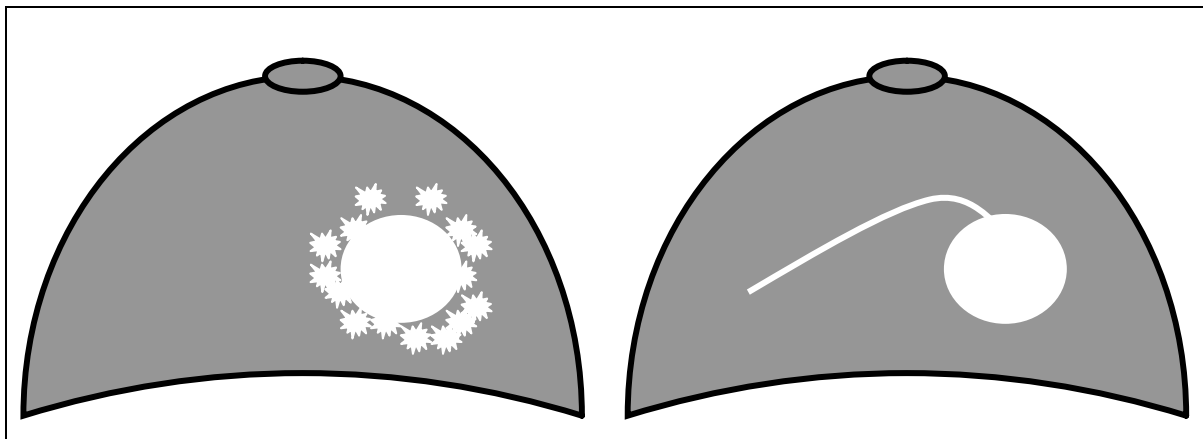


Abbildung 4.4: Perifokales Ödem und zuführendes Gefäß

Schematische Darstellung eines perifokalen Ödems (links) und eines zuführenden Gefäßes (rechts)

4.5 Histologische Aufarbeitung

Bei in der MR-Mammographie als suspekt eingestuften Läsionen wurde zu einer histologischen Abklärung derselben geraten. Das suspekte Areal wurde entweder durch Probeexzision entfernt, oder es erfolgte sonographisch-, stereotaktisch- oder MR-gestützt eine Stanzbiopsie. Bei unklarer Lokalisation erfolgte vor der Intervention eine Drahtmarkierung

und bei der Aufarbeitung mikrokalkenthaltenden Gewebes zwecks Kontrolle eine Präparate-Radiographie. Die histologische Aufarbeitung erfolgte am Institut für Pathologie des UKJ jeweils durch in der Pathologie der Mamma erfahrene Pathologen.

4.6 Statistische Methoden

Ein Großteil der statistischen Verfahren wurde mit der Statistikanwendung PASW Statistics 18.0 (Chicago, Illinois, USA) durchgeführt. Zur Überprüfung der einzelnen Merkmale auf signifikante Unterschiede im Auftreten bei benignen und malignen Anreicherungen fanden der Chi-Quadrat-Test und der exakte Test nach Fisher Anwendung. Letzterer wurde durchgeführt, falls eine der erwarteten Häufigkeiten weniger als fünf betrug. Gemäß der Nullhypothese H_0 unterscheiden sich die Gruppen anhand des untersuchten Merkmals nicht. Die Alternativhypothese H_1 besagt, dass ein Unterschied unabhängig von der Richtung besteht. Es wurde zweiseitig mit einem Signifikanzniveau von $\alpha=0,05$ getestet. Die Nullhypothese H_0 wurde bei P-Werten von $\alpha \leq 0,05$ zugunsten der Alternativhypothese H_1 verworfen. Ein Unterschied zwischen den Dignitäten hinsichtlich des Auftretens eines Merkmals war somit bei $P \leq 0,05$ statistisch signifikant, bei $P \leq 0,01$ hochsignifikant und bei $P \leq 0,001$ höchstsignifikant.

Mit dem Verfahren der binären logistischen Regression wurde in Abhängigkeit der Ausprägung einzelner MR-mammographisch sichtbarer Merkmale die Wahrscheinlichkeit für das Vorliegen eines Malignoms berechnet. Dies geschah zunächst unter Einschluss aller analysierten Zeichen. Es erfolgte aufgrund des kategorialen Charakters der Variablen eine Umwandlung in Dummy-Variablen, als Kontrast wurde die Einstellung *Indikator* gewählt. Danach fand durch Nutzung von Vorwärts- und Rückwärtsselektionsalgorithmen eine Eliminierung von Prädiktorvariablen, die das Regressionsmodell nicht signifikant verbessern, statt. Im Anschluss wurde ein optimiertes Regressionsmodell mit lediglich fünf Deskriptorvariablen erstellt. Die Auswahl der Kriterien stützte sich auf den Vorwärtsselektionsalgorithmus, wobei die ersten fünf ins Modell aufgenommenen MR-Zeichen verwendet wurden. Dieses neu gewonnene Konstrukt kam zum Vergleich mit einem Modell, welches die Malignomwahrscheinlichkeit anhand der fünf etablierten Deskriptoren *Form*, *Rand*, *Septen*, *zeitliches Signalintensitätsverhalten* und *KM-Anreicherungsmuster* errechnete.

Mit Hilfe von ROC (receiver operating characteristic)-Kurven gelang die Schätzung der Modellgüte. In der Kurve wird für jede in der binär-logistischen Regression errechnete Malignitätswahrscheinlichkeit graphisch dargestellt, wie hoch die Sensitivität und Spezifität wären unter der Annahme, dass der entsprechende Wahrscheinlichkeitswert als Cut-off value zwischen Gut- und Bösartigkeit dienen würde. Die Güte und Vorhersagequalität des entsprechenden Regressionsmodells wird hierbei durch die Fläche (area under curve = AUC-Wert) unter der ROC-Kurve quantifiziert.

Befunde, deren Malignomwahrscheinlichkeit mindestens 95% betrugen, betrachtete man als sicher maligne, solche mit einer Malignitätswahrscheinlichkeit kleiner gleich 5% als sicher benigne. Es wurde jeweils ausgewertet, wie viele falsch-positive sich unter den sicher bösartigen Herden bzw. wie viele falsch-negative als sicher benigne kategorisiert wurden.

Mit der Software MedCalc Version 11.1 (Mariakerke, Belgien) konnte man durch Implementierung der Regressionsmodelle prüfen, ob sich die aus ihnen ermittelten ROC-Kurven signifikant unterschieden.

5 Ergebnisse

5.1 Häufigkeit von MR-Zeichen

Die Tabellen 5.1, 5.2 und 5.3 geben die Häufigkeiten der morphologischen und kinetischen Deskriptoren sowie die Prävalenz der Zeichen des den suspekten Befund umgebenden Mammaparenchyms wieder. Die P-Werte geben Auskunft darüber, ob sich das betrachtete Merkmal im Auftreten bei malignen und benignen Befunden signifikant unterscheidet.

Tabelle 5.1: Häufigkeiten morphologischer Zeichen

MR-Zeichen und dessen Ausprägung	Absolute Häufigkeiten innerhalb der Malignome	Prozentuale Verteilung innerhalb der Malignome	Absolute Häufigkeiten innerhalb der benignen Befunde	Prozentuale Verteilung innerhalb der benignen Befunde	P-Wert
Form					
rund	233	36% (233/648)	235	54% (235/436)	<0,001
lobuliert	35	5% (35/648)	64	15% (64/436)	<0,001
irregulär	380	59% (380/648)	137	31% (137/436)	<0,001
Rand					
scharf	79	12% (79/648)	194	44% (194/436)	<0,001
unscharf	569	88% (569/648)	242	56% (242/436)	<0,001
Septen					
keine	630	97% (630/648)	361	83% (361/436)	<0,001
vorhanden	10	2% (10/648)	59	13% (59/436)	<0,001
anreichernd	8	1% (8/648)	16	4% (16/436)	0,008
T2-Helligkeit					
iso/dunkel	628	97% (628/648)	342	78% (342/436)	<0,001
hell	20	3% (20/648)	94	22% (94/436)	<0,001
T1-Binnen- struktur					
homogen	213	33% (213/648)	210	48% (210/436)	<0,001
inhomogen	435	67% (435/648)	226	52% (226/436)	<0,001
T1-Helligkeit					
hell	7	1% (7/648)	20	4% (20/436)	<0,001
iso	383	59% (383/648)	282	65% (282/436)	0,065
dunkel	258	40% (258/648)	134	31% (134/436)	0,002
Hook- Zeichen					
negativ	458	71% (458/648)	422	97% (422/436)	<0,001
positiv	190	29% (190/648)	14	3% (14/436)	<0,001
Root-Zeichen					
negativ	172	27% (172/648)	349	80% (349/436)	<0,001
positiv	476	73% (476/648)	87	20% (87/436)	<0,001

Tabelle 5.2: Häufigkeiten KM-dynamischer Zeichen

MR-Zeichen und dessen Ausprägung	Absolute Häufigkeiten innerhalb der Malignome	Prozentuale Verteilung innerhalb der Malignome	Absolute Häufigkeiten innerhalb der benignen Befunde	Prozentuale Verteilung innerhalb der benignen Befunde	P-Wert
Postinitiale					
KM-Dynamik					
weiterer Anstieg	64	10% (64/648)	234	54% (234/436)	<0,001
Plateau	151	23% (151/648)	109	25% (109/436)	0,521
wash-out	433	67% (433/648)	93	21% (93/436)	<0,001
Anreicherungs-					
verhalten					
homogen	81	13% (81/648)	153	35% (153/436)	<0,001
inhomogen	349	54% (349/648)	240	55% (240/436)	0,041
zentrifugal	9	1% (9/648)	14	3% (14/436)	0,700
zentripetal	209	32% (209/648)	29	7% (29/436)	<0,001
Blooming					
negativ	272	42% (272/648)	374	86% (374/436)	<0,001
positiv	376	58% (376/648)	62	14% (62/436)	<0,001

Tabelle 5.3: Häufigkeiten der die Läsionsumgebung beschreibenden Zeichen

MR-Zeichen und dessen Ausprägung	Absolute Häufigkeiten innerhalb der Malignome	Prozentuale Verteilung innerhalb der Malignome	Absolute Häufigkeiten innerhalb der benignen Befunde	Prozentuale Verteilung innerhalb der benignen Befunde	P-Wert
Unilaterales Ödem					
kein Ödem	364	56% (364/648)	392	90% (392/436)	<0,001
perifokal diffus	245	38% (245/648)	29	7% (29/436)	<0,001
ipsilateral	39	6% (39/648)	15	3% (15/436)	0,056
Gefäß					
unauffällig	251	39% (251/648)	324	74% (324/436)	<0,001
prominent	91	14% (91/648)	58	13% (58/436)	0,728
zuführend	306	47% (306/648)	54	13% (54/436)	<0,001
Cutisver- dickung					
negativ	574	89% (574/648)	420	96% (420/436)	<0,001
positiv	74	11% (74/648)	16	4% (16/436)	<0,001
Mamillen- saumunter- brechung					
negativ	563	87% (563/648)	407	93% (407/436)	0,001
positiv	85	13% (85/648)	29	7% (29/436)	0,001

5.2 Binäre logistische Regressionsmodelle

5.2.1 Einschlussmethode, Vorwärts- und Rückwärtsselektion

Die Einschlussmethode mit Aufnahme aller 15 evaluierten Variablen ergab eine zu 83,8% korrekte Differenzierung zwischen malignen und benignen Veränderungen. Bei der stufenweisen Aufnahme der unabhängigen Variablen wurde eine um 0,3% geringere Genauigkeit von 83,5% erreicht. Zur Berechnung für das Vorliegen eines malignen Befundes wurden der Reihenfolge nach folgende Deskriptoren ins Regressionsmodell aufgenommen: *Root-Zeichen*, *postinitiale Kinetik*, *räumliches KM-Anreicherungsmuster*, *T2-Helligkeit*, *Ödem*, *Form*, *Hook-Zeichen*, *Blooming-Zeichen*, *Septen*, *T1-Helligkeit*. Informationen zu *Berandung*, *Cutisverdickung*, *Mamillensaumunterbrechung*, *T1-Binnenstruktur* oder *Gefäßauffälligkeiten* fanden keine Berücksichtigung.

Das über Rückwärtsselektion errechnete Regressionsmodell, bei welchem erst alle unabhängigen Variablen Verwendung fanden und dann schrittweise die Merkmale *Mamillensaumunterbrechung*, *Cutisverdickung* und *T1-Binnenstruktur* eliminiert wurden, kam auf eine ebenfalls zu 83,5% korrekte Klassifikation der Befunde.

Die AUC ergab jeweils 0,91. Die drei ROC-Kurven für die jeweiligen regressionsmodellabhängigen Malignitätswahrscheinlichkeiten unterschieden sich mit einem P von jeweils $>0,05$ nicht signifikant. Tab. 5.4 gibt einen zusammenfassenden Überblick über die drei hier erörterten Modelle mitsamt deren statistischen Kennwerten, Tab. 5.5 zeigt den Vergleich der Modelle hinsichtlich deren AUC und signifikanter Unterschiede.

Tabelle 5.4: Statistische Kennzahlen der initial errechneten Regressionsmodelle

Methode	Anteil richtig Klassifizierter	AUC	Sensitivität	Spezifität	PPV	NPV
Einschluss- methode	83,8% (908/1084)	0,911	86,6% (561/648)	79,6% (347/436)	86,3% (561/650)	80,0% (347/434)
Vorwärts- selektion	83,5% (905/1084)	0,909	86,7% (562/648)	78,7% (343/436)	85,8% (562/655)	80,0% (343/429)
Rückwärts- selektion	83,5% (905/1084)	0,910	86,3% (559/648)	79,4% (346/436)	86,1% (559/649)	79,5% (346/435)

Tabelle 5.5: AUC-Vergleich der initial errechneten Modelle

Zu vergleichende Modelle	AUC der Modelle	P-Wert
Einschluss vs. Vorwärts	0,911 vs. 0,909	0,458
Einschluss vs. Rückwärts	0,911 vs. 0,910	0,391
Vorwärts vs. Rückwärts	0,909 vs. 0,910	0,722

5.2.2 Binäre logistische Modelle mit fünf Deskriptoren

Entsprechend der ersten fünf durch die Vorwärtsselektion ausgewählten Deskriptoren wurde ein Modell errechnet, welches *Root-Zeichen*, *zeitliches KM-Anreicherungsmuster*, *räumliches KM-Anreicherungsmuster*, *T2-Signalintensität* und das Vorhandensein *ipsilateralen Ödeme* zur Malignitätsvorhersage berücksichtigt. Tab. 5.6 stellt die errechneten statistischen Kennwerte des Modells dar und zeigt den Vergleich zu einem fünf Deskriptoren enthaltenden Modell mit den Eigenschaften *Form*, *Rand*, *Septen*, *zeitliches KM-Anreicherungsmuster* und *KM-Anreicherungsverhalten*.

Ersteres Modell wies eine diagnostische Genauigkeit von 81,5% (884/1084) gegenüber dem Modell mit den etablierten Parametern auf, welches zu 80,4% (871/1084) korrekt klassifizierte. Die AUC war bei ersterem Modell mit 0,894 signifikant höher als die AUC von 0,875 bei letzterem (P=0,017). Wie in Tab. 5.7 deutlich wird, unterschieden sich die AUC der beiden Fünf-Deskriptoren-Modelle hoch bis höchst signifikant von denen der drei bereits erwähnten Regressionsmodelle. Die Tabellen 5.8 und 5.9 geben die statistischen

Modellparameter der Fünf-Deskriptorenmodelle zur Errechnung der Wahrscheinlichkeit wieder. Diese Berechnung der Wahrscheinlichkeit für Malignität ist im Anhang in Kapitel 9.1.1 erläutert.

Tabelle 5.6: Statistische Kennwerte zu den Modellen mit fünf Deskriptoren

Methode	Anteil richtig Klassifizierter	AUC	Sensitivität	Spezifität	PPV	NPV
5 Prädiktoren *	81,5 % (884/1084)	0,894	82,7 % (536/648)	79,8 % (348/436)	84,5% (536/634)	75,7% (348/460)
5 Prädiktoren **	80,4 % (871/1084)	0,875	84,6 % (548/648)	74,1 % (323/436)	82,9 % (548/661)	76,4 % (323/423)

* Root-Zeichen, postinitiale Dynamik, internes Anreicherungsmuster, T2-Signalintensität, Ödem

** Form, Rand, Septen, postinitiale Dynamik, KM-Anreicherungsverhalten

Tabelle 5.7: AUC-Vergleich initial Erstellter mit den Fünf-Deskriptoren-Modellen

Zu vergleichende Modelle	AUC der Modelle	P-Wert
Einschluss vs. 5 Prädiktoren *	0,911 vs. 0,894	0,001
Einschluss vs. 5 Prädiktoren **	0,911 vs. 0,875	<0,001
Vorwärts vs. 5 Prädiktoren *	0,909 vs. 0,894	0,002
Vorwärts vs. 5 Prädiktoren **	0,909 vs. 0,875	<0,001
Rückwärts vs. 5 Prädiktoren *	0,910 vs. 0,894	0,002
Rückwärts vs. 5 Prädiktoren **	0,910 vs. 0,875	<0,001
5 Prädiktoren * vs. 5 Prädiktoren **	0,894 vs. 0,875	0,017

* Root-Zeichen, postinitiale Dynamik, internes Anreicherungsmuster, T2-Signalintensität, Ödem

** Form, Rand, Septen, postinitiale Dynamik, KM-Anreicherungsverhalten

Tabelle 5.8: Optimiertes fünf Deskriptoren enthaltendes Regressionsmodell

Deskriptorvariable	Ausprägung	Odds Ratio	KI 95%	Signifikanz	Koeffizient (ln OR) für Wahrscheinlichkeitsberechnung
Internes Anreicherungs-muster	zentrifugal	0,914	[0,300; 2,785]	0,875	-0,090
	heterogen	2,642	[1,763; 3,959]	<0,001	0,971
	zentripetal	5,796	[3,205; 10,482]	<0,001	1,757
T2-Helligkeit	hyperintens	0,208	[0,113; 0,381]	<0,001	-1,571
Root-Zeichen	positiv	4,163	[2,934; 5,906]	<0,001	1,426
Ödem	diffus	2,032	[0,910; 4,537]	0,084	0,709
	ipsilateral				
	perifokal	3,447	[2,062; 5,763]	<0,001	1,238
Kurvenverlauf	Plateau	4,156	[2,675; 6,458]	<0,001	1,425
	wash-out	9,727	[6,383; 14,825]	<0,001	2,275

Konstante= -0,997 , nicht enthaltene Punktwerte für bestimmte Ausprägungen entsprechen einem Wert von Null

Tabelle 5.9: Fünf-Deskriptoren-Regressionsmodell mit konventionellen Zeichen

Deskriptorvariable	Ausprägung	Odds Ratio	KI 95%	Signifikanz	Koeffizient (ln OR) für Wahrscheinlichkeits- berechnung
Form	lobuliert	0,707	[0,378; 1,322]	0,277	-0,347
	irregulär	2,100	[1,446; 3,052]	<0,001	0,742
Rand	unscharf	2,540	[1,690; 3,819]	<0,001	0,932
Internes Anreicherungsmuster	zentrifugal	1,072	[0,353; 3,254]	0,903	0,069
	heterogen	2,663	[1,763; 4,023]	<0,001	0,979
	zentripetal	7,975	[4,557; 13,955]	<0,001	2,076
Septen	vorhanden	0,234	[0,105; 0,521]	<0,001	-1,454
	anreichernd	0,233	[0,077; 0,705]	0,01	-1,459
Kurvenverlauf	Plateau	5,788	[3,775; 8,874]	<0,001	1,756
	wash-out	18,278	[12,018; 27,799]	<0,001	2,906

Konst = -3,156 , nicht enthaltene Punktwerte für bestimmte Ausprägungen entsprechen einem Wert von Null

5.3 Erfassung von Läsionen mit sicherer Klassifizierbarkeit

5.3.1 Erfassung sicher maligner Läsionen

Durch ein Regressionsmodell allein lassen sich maximal 235 Befunde bei 648 Malignomen mit mehr als 95%iger Wahrscheinlichkeit als bösartig zuordnen. Dies wäre möglich, wenn man das Regressionsmodell unter Einschluss aller evaluierten Deskriptoren nutzt. Mit 229 bzw. 234 Fällen ist die Anzahl der detektierten Fälle bei den Vorwärts- bzw. Rückwärtsselektionsmodellen vergleichbar hoch. Da modellabhängig drei bzw. vier Läsionen fälschlicherweise detektiert wurden, betrug der Anteil an falsch-positiven Fällen unter den als sicher maligne Klassifizierten 1,3 bis 1,7%. Unter Nutzung des Modells mit etablierten MR-Zeichen wurden 98 Raumforderungen als sicher maligne bewertet, die Falsch-Positiv-Rate war hierbei mit 2% (2/98) am höchsten. Das andere Fünf-Faktoren-Modell ermittelte 176 sicher maligne Befunde, davon zwei Falsch-Positive. Somit lag hier mit 1,1% die niedrigste Falsch-Positiv-Rate vor.

In relativen Häufigkeiten ausgedrückt gelang somit im günstigsten Fall die Identifikation von 35,6% der Brustkrebsfälle mit an Sicherheit grenzender Wahrscheinlichkeit, wenn man sämtliche evaluierte Deskriptoren nutzte. Durch Beschränkung auf fünf Deskriptoren mit teils neuen verwendeten Zeichen wurden maximal 26,9% erkannt, unter Anwendung der etablierten fünf Kriterien 14,8% (vgl. Tab. 5.10). Ein Beispiel für einen sicher malignen Befund ist in Abb. 5.1 dargestellt.

Tabelle 5.10: Modellabhängige Erfassung sicher maligner Befunde

Regressionsmodell	Anzahl sicher maligne detektierter Läsionen	Anzahl falsch- positiver Befunde	Anteil der richtig- positiv als sicher maligne Klassifizier- ten an allen histo- logisch verifizierten Malignomen	Falsch-Positiv-Rate (Falsch-Positive/ Sicher-Maligne)
Einschluss-Meth.	235	4	35,6% (231/648)	1,7% (4/235)
Vorwärts-Meth.	229	4	34,7% (225/648)	1,7% (4/229)
Rückwärts-Meth.	234	3	35,6% (231/648)	1,3% (3/234)
5 Deskriptoren [*]	176	2	26,9% (174/648)	1,1% (2/176)
5 Deskriptoren ^{**}	98	2	14,8% (96/648)	2,0% (2/98)

^{*} Root-Zeichen, postinitiale Dynamik, internes Anreicherungsmuster, T2-Signalintensität, Ödem

^{**} Form, Rand, Septen, postinitiale Dynamik, KM-Anreicherungsverhalten

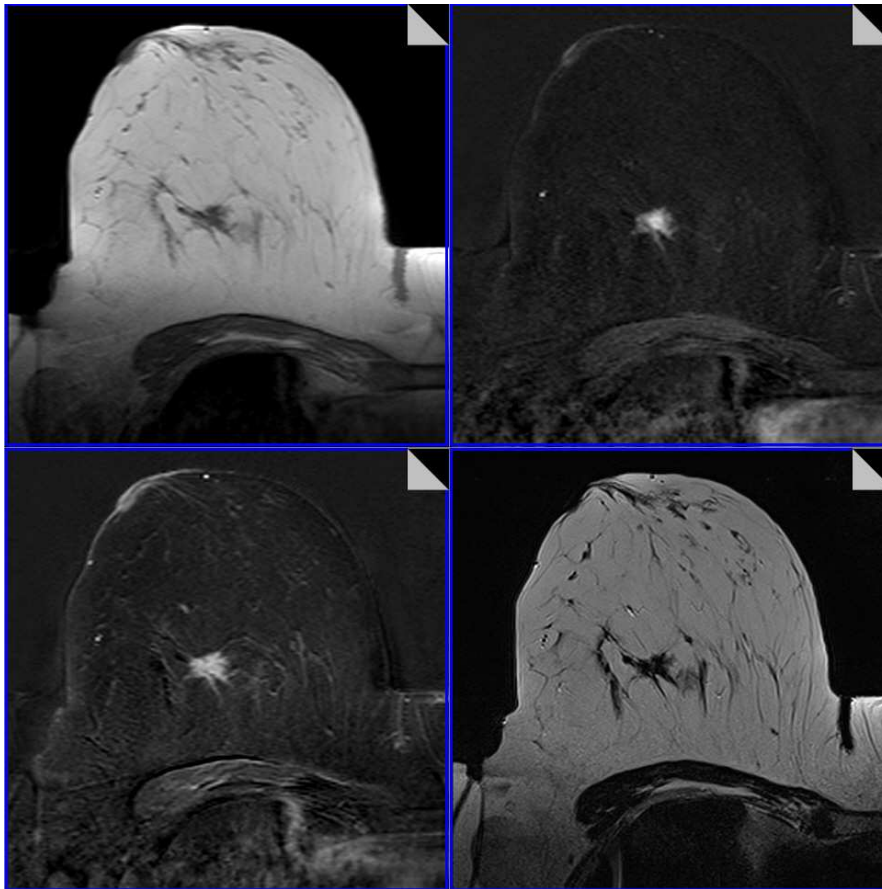


Abbildung 5.1: Sicher maligner Befund

Deskriptorausprägungen: Root-sign positiv, wash-out, inhomogen, hypointens in T2, kein Ödem

Bilderfolge: T1w nativ (links oben), Subtraktionsaufnahme 1 min nach KM-Applikation (rechts oben),
Subtraktionsaufnahme 7 min nach KM-Applikation (links unten), T2-TSE (rechts unten)

5.3.2 Erfassung sicher benigner Läsionen

Mittels Durchführung der binären logistischen Regression konnten unter Einschluss aller Deskriptoren 73 Läsionen mit einer maximal 5%igen Wahrscheinlichkeit für Malignität ermittelt werden, 68 davon waren korrekt erfasst. Unter Berücksichtigung der fünf etablierten Zeichen wurden 69 Anreicherungen auf diese Art und Weise detektiert, in 2,9% (2/69) war die Klassifizierung falsch, was bei Einschluss aller Deskriptoren zu 6,8% (5/73) der Fall war. Das andere Fünf-Faktoren-Modell ermittelte 43 benigne Fälle, 3 davon waren bösartig, was eine Falsch-Negativ-Rate von 7% (3/43) ergab. Hierbei wurden 9,2% (40/436) der gutartigen Läsionen korrekt klassifiziert, bei dem anderen Fünf-Deskriptoren-Modell 15,4% (67/436). 15,6% (68/436) wurden unter Nutzung sämtlicher Informationen ermittelt (vgl. Tab. 5.11). Exemplarisch findet sich eine nach dem Fünf-Deskriptoren-Modell als sicher gutartig zu klassifizierende Läsion unter Anwendung etablierter häufig publizierter Zeichen in Abb. 5.2.

Tabelle 5.11: Modellabhängige Erfassung sicher benigner Befunde

Regressionsmodell	Anzahl sicher benigne detektierter Läsionen	Anzahl falsch- negativer Befunde	Anteil der richtig- negativ als sicher benigne Klassifizierten an allen histologisch verifizierten benignen Befunden	Falsch-Negativ- Rate (Falsch- Negative/Sicher- Benigne)
Einschluss-Meth.	73	5	15,6% (68/436)	6,8% (5/73)
Vorwärts-Meth.	59	5	12,4% (54/436)	8,5% (5/59)
Rückwärts-Meth.	71	5	15,1% (66/436)	7,0% (5/71)
5 Deskriptoren*	43	3	9,2% (40/436)	7,0% (3/43)
5 Deskriptoren**	69	2	15,4% (67/436)	2,9% (2/69)

* Root-Zeichen, postinitiale Dynamik, internes Anreicherungsmuster, T2-Signalintensität, Ödem

** Form, Rand, Septen, postinitiale Dynamik, KM-Anreicherungsverhalten

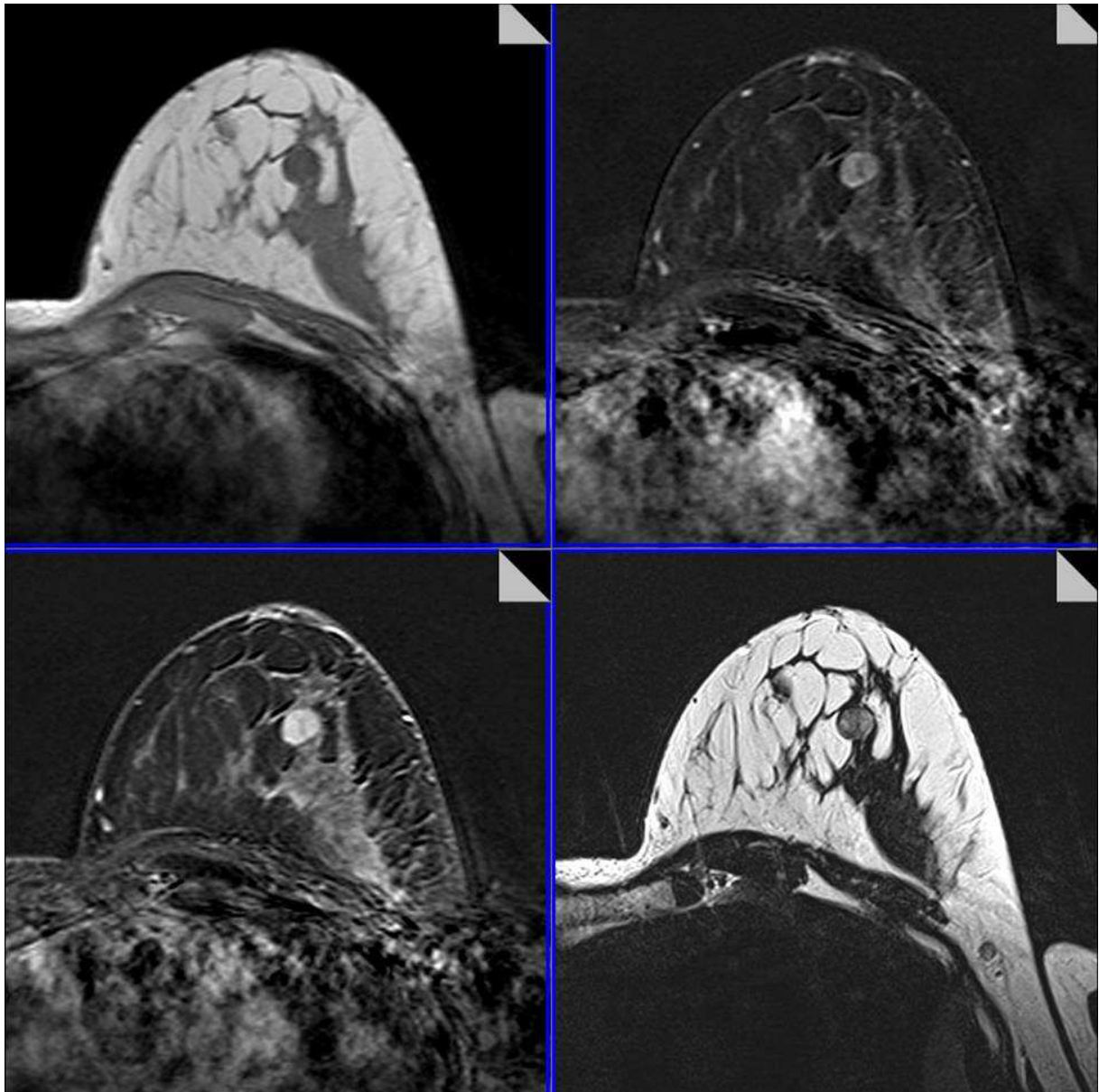


Abbildung 5.2: Sicher gutartiger Befund

Deskriptorenzuweisung: Form oval, glatt berandet, dunkle Septen, persistierende KM-Anreicherung mit homogenem Anreicherungsmuster

Bilderfolge: T1w nativ (links oben), Subtraktionsaufnahme 1 min nach KM-Applikation (rechts oben), Subtraktionsaufnahme 7 min nach KM-Applikation (links unten), T2-TSE (rechts unten)

5.3.3 Summarische Erfassung sicher maligner und benigner Läsionen

Unter Einschluss sämtlicher Deskriptoren gelang zu 27,6% eine sichere Klassifikation benigner und maligner Befunde. Die Limitation auf fünf Deskriptorvariablen ist mit 19,7%

bzw. 15,0% weniger erfolgreich, das Modell mit den fünf konventionellen Zeichen schneidet insgesamt am schlechtesten ab. Der relative Anteil der falsch klassifizierten Läsionen bewegt sich zwischen den Modellen auf vergleichbarem Niveau, bei Nutzung des durch diese Arbeit erstellten optimierten Fünf-Deskriptoren-Modells ist er am geringsten (vgl. Tab. 5.12)

Tabelle 5.12: Modellabhängige Erfassung sicher maligner und benigner Befunde

Regressionsmodell	Anzahl der sicher klassifizierten Läsionen	Anzahl falscher Befunde	Anteil der histologisch verifizierten Läsionen, der mit hoher Sicherheit richtig klassi- fiziert wurde	Anteil Falscher an den Läsionen, die sicher klassifiziert wurden
Einschluss	308	9	27,6% (299/1084)	2,9% (9/308)
Vorwärts	288	9	25,7% (279/1084)	3,1% (9/288)
Rückwärts	305	8	27,4% (297/1084)	2,6% (8/305)
5 Deskriptoren [*]	219	5	19,7% (214/1084)	2,3% (5/219)
5 Deskriptoren ^{**}	167	4	15,0% (163/1084)	2,4% (4/167)

* Root-Zeichen, postinitiale Dynamik, internes Anreicherungsmuster, T2-Signalintensität, Ödem

** Form, Rand, Septen, postinitiale Dynamik, KM-Anreicherungsverhalten

5.3.4 Zeichenkombinationen bei sicher klassifizierbaren Läsionen

Eine Wahrscheinlichkeit von mindestens 95% für Malignität ergab sich ab einer Summe aus modellspezifischer Konstante und zutreffenden Koeffizienten von 2,944. Zur Detektion solcher Befunde wurde das Modell mit den fünf Deskriptoren *Root-sign*, *räumliches und zeitliches KM-Anreicherungsmuster*, *T2-Helligkeit* und *Ödem* verwendet. Es konnte bei diesem Modell ein Score von maximal 5,699 erreicht werden. Tab. 5.13 zeigt die möglichen Merkmalskonstellationen. Die Vorgehensweise der Herleitung möglicher Merkmalskombinationen wird detailliert im Anhang dieser Arbeit (Kapitel 9.1.2) beschrieben. Analog dazu sind in Tab. 5.14 die Merkmalskombinationen für das Vorliegen eines wahrscheinlich gutartigen Befundes angegeben. Aufgrund der im Vergleich zu den anderen Modellen dargebotenen geringen Falsch-Negativ-Rate wurde das Fünf-Deskriptoren-Modell mit den konventionellen MR-Zeichen herangezogen.

Tabelle 5.13: Mögliche Merkmalskombinationen für sicher maligne Befunde

Zeitliche Konstanz des KM-Anreicherungsverhaltens	Root-Zeichen	Räumliches KM-Anreicherungsverhalten	Ödem	T2-Helligkeit	Score
wash-out	positiv	zentripetal	perifokal	isointens/dunkel	5,699
wash-out	positiv	zentripetal	diff. ips.	isointens/dunkel	5,197
wash-out	positiv	zentripetal	negativ	isointens/dunkel	4,461
wash-out	positiv	homogen	perifokal	isointens/dunkel	3,942
wash-out	positiv	zentrifugal	perifokal	isointens/dunkel	3,852
wash-out	positiv	heterogen	perifokal	isointens/dunkel	4,913
wash-out	negativ	zentripetal	perifokal	isointens/dunkel	4,273
wash-out	positiv	zentripetal	perifokal	hell	4,128
wash-out	negativ	heterogen	perifokal	isointens/dunkel	3,387
wash-out	positiv	heterogen	perifokal	hell	3,342
wash-out	positiv	homogen	diff. ips.	isointens/dunkel	3,413
wash-out	positiv	zentrifugal	diff. ips.	isointens/dunkel	3,323
wash-out	positiv	heterogen	diff. ips.	isointens/dunkel	4,384
wash-out	positiv	zentripetal	diff. ips.	hell	3,599
wash-out	negativ	heterogen	diff. ips.	isointens/dunkel	2,958
wash-out	negativ	zentripetal	diff. ips.	isointens/dunkel	3,744
wash-out	positiv	heterogen	negativ	isointens/dunkel	3,675
wash-out	negativ	zentripetal	negativ	isointens/dunkel	3,035
Plateau	positiv	heterogen	diff. ips.	isointens/dunkel	3,534
Plateau	positiv	homogen	perifokal	isointens/dunkel	3,092
Plateau	positiv	zentrifugal	perifokal	isointens/dunkel	3,002
Plateau	positiv	zentripetal	perifokal	hell	3,278
Plateau	negativ	zentripetal	perifokal	isointens/dunkel	3,423
Plateau	positiv	zentripetal	negativ	isointens/dunkel	3,611
Plateau	positiv	zentripetal	diff. ips.	isointens/dunkel	4,320
Plateau	positiv	heterogen	perifokal	isointens/dunkel	4,063
Plateau	positiv	zentripetal	perifokal	isointens/dunkel	4,849
persistierend	positiv	zentripetal	perifokal	isointens/dunkel	3,424

Tabelle 5.14: Mögliche Merkmalskombinationen für sicher benigne Befunde

Zeitliche Konstanz des KM-Anreicherungs- verhaltens	Räumliches KM- Anreicherungs- muster	Rand	Form	Septen	Score
Plateau	homogen	scharf	lobuliert	vorhanden	-3,201
Plateau	homogen	scharf	lobuliert	anreichernd	-3,206
kontinuierlich	heterogen	scharf	rund/oval	vorhanden	-3,631
kontinuierlich	heterogen	scharf	rund/oval	anreichernd	-3,636
kontinuierlich	heterogen	scharf	lobuliert	vorhanden	-3,978
kontinuierlich	heterogen	scharf	lobuliert	anreichernd	-3,983
kontinuierlich	zentrifugal	scharf	rund/oval	vorhanden	-4,541
kontinuierlich	zentrifugal	scharf	rund/oval	anreichernd	-4,546
kontinuierlich	zentrifugal	scharf	lobuliert	vorhanden	-4,888
kontinuierlich	zentrifugal	scharf	lobuliert	anreichernd	-4,893
kontinuierlich	zentrifugal	unscharf	rund/oval	vorhanden	-4,541
kontinuierlich	zentrifugal	unscharf	rund/oval	anreichernd	-4,546
kontinuierlich	zentrifugal	scharf	irregulär	vorhanden	-3,799
kontinuierlich	zentrifugal	scharf	irregulär	anreichernd	-3,804
kontinuierlich	zentrifugal	unscharf	lobuliert	vorhanden	-4,888
kontinuierlich	zentrifugal	unscharf	lobuliert	anreichernd	-4,893
kontinuierlich	homogen	scharf	rund/oval	vorhanden	-4,610
kontinuierlich	homogen	scharf	rund/oval	anreichernd	-4,615
kontinuierlich	homogen	scharf	lobuliert	vorhanden	-4,957
kontinuierlich	homogen	scharf	lobuliert	anreichernd	-4,962
kontinuierlich	homogen	unscharf	rund/oval	vorhanden	-3,678
kontinuierlich	homogen	unscharf	rund/oval	anreichernd	-3,683
kontinuierlich	homogen	scharf	irregulär	vorhanden	-3,868
kontinuierlich	homogen	scharf	irregulär	anreichernd	-3,873
kontinuierlich	homogen	unscharf	lobuliert	vorhanden	-4,957
kontinuierlich	homogen	unscharf	lobuliert	anreichernd	-4,962
kontinuierlich	homogen	scharf	rund/oval	keine	-3,156
kontinuierlich	zentrifugal	scharf	rund/oval	keine	-3,087
kontinuierlich	homogen	scharf	lobuliert	keine	-3,503
kontinuierlich	zentrifugal	scharf	lobuliert	keine	-3,434

6 Diskussion

6.1 Interpretationsmodelle in der MRM-Diagnostik

Die kontrastmittelunterstützte kernspintomographische Untersuchung der Mammæ gilt als hochsensitives Verfahren zur Brustkrebserkennung. Nach Veröffentlichung von Studien zur MRM mit hohen Sensitivitäten begleitet von sich weniger gut ausnehmender Spezifität wurde die MR-Diagnostik dahingehend verbessert, anhand bestimmter Zeichen in den MR-Bildern eine bessere Diskriminierung zwischen Gut- und Bösartigkeit eines Tumors zu ermöglichen.

Die hohe Sensitivität der MRM rührt daher, dass bösartige Tumoren ab einer Größe von 3 mm potentiell dazu in der Lage sind, KM anzureichern (Kaiser 1989). Ein gewisser Anteil falsch-positiver Befunde resultiert daher, dass jede Anreicherung zunächst als malignomverdächtig betrachtet wird, es sich aber auch um einen gutartigen Befund handeln kann. Abhängig von der zugrunde liegenden Studienpopulation existieren deswegen Daten zur Spezifität zwischen 37%-97% (Harms et al. 1993a, Kaiser 1993). Die kontrovers diskutierte, teils als niedrig angegebene Spezifität ist ein Grund für die nur langsam voranschreitende Etablierung der MRM in der klinischen Routine. Zwecks besserer Differenzierung wurden zahlreiche für die Befunderhebung der MRM spezifische Zeichen beschrieben. Bei der Erforschung solcher Kriterien ist dabei zwischen Merkmalen zu unterscheiden, die entweder architektonisch-struktureller Art sind oder die aufgrund des Verhaltens der Kontrastmittelkinetik bestimmte Schlussfolgerungen zulassen. Eine kombinierte Berücksichtigung morphologischer und kinetischer Zeichen liefert bessere Ergebnisse, als wenn lediglich die morphologischen oder dynamischen Charakteristika jeweils für sich allein untersucht würden (Schnall et al. 2001). Nachdem sich vor allem die Betrachtung des zeitlichen und räumlichen Verlaufs des Kontrastmittelanreicherungsverhaltens, der Herdberandung, der Herdform und die Prüfung auf Vorhandensein von Herdseptierungen etabliert hatte, führten Malich und Kollegen zahlreiche weitere Deskriptoren ein, die insbesondere bei schwierigen Fällen die diagnostische Zuordnung zu erleichtern scheinen (Fischer et al. 2005, Malich et al. 2005).

Die wohl umfassendste deskriptive Publikation zu MR-mammographischen Zeichen erfolgte durch Kaiser, in welcher deren 147 beschrieben sind. Diese dienen jedoch nicht allein der Dignitätsdiagnostik einer suspekten Läsion, es wird u. a. auch auf Brustdarstellungen unter Chemotherapie, bei Z. n. Radiatio oder bei Z. n. Implantatseinbringung eingegangen (Kaiser

2007). Durch Zuordnung gewisser definierter Charakteristika zu bestimmten Läsionseigenschaften verfügt der Diagnostiker somit aktuell über eine Vielzahl an richtungsweisenden Läsionsmerkmalen, anhand derer er seinen Befund und die sich daraus ergebenden Konsequenzen abwägen kann.

Nach wie vor erfordert die MR-Mammographie reichlich Erfahrung des Untersuchers, und die diagnostische Festlegung fällt oft schwer. Arbeiten zur Röntgenmammographie kamen zu dem Ergebnis, dass Entscheidungshilfen einen Beitrag zur Verbesserung der diagnostischen Genauigkeit leisten können (Swets et al. 1991, Wu et al. 1993). Im Rahmen der vorliegenden Arbeit wurden mehrere verschiedene Interpretationsmodelle erstellt, bei denen jeweils unterschiedliche Merkmalsausprägungen aufgenommen wurden. Da die Interpretation bei bildgebenden Verfahren stark von der Erfahrung des Befunders abhängt, wurde das Ziel der Modellerstellung vor allem in der Intention verfolgt, speziell dem unerfahrenen Untersucher eine Entscheidungshilfe zu bieten.

Um unter Berücksichtigung bestimmter Deskriptoren eine Differenzierung hinsichtlich der Dignität einer in der MRM anreichernden Läsion vorzunehmen, ist es unverzichtbar, diverse Läsionseigenschaften zu bewerten und aufgrund deren konzertierter Betrachtung den Befund abzuwägen. Die Notwendigkeit ergibt sich daraus, dass die Bewertung eines Zeichens allein keine zufriedenstellend hochwertige Differenzierung zulässt. Es existieren zwar Merkmale, bei deren Vorliegen man sich mit hoher Sicherheit auf einen Befund festlegen kann, eine annähernd 100%ige Sicherheit lässt sich hierdurch nicht erreichen, und weisen solche Zeichen einen hohen prädiktiven Wert auf, so ist deren Prävalenz mäßig hoch, was sich folglich auch mit einer niedrigen Sensitivität bemerkbar macht. Andererseits existieren Merkmale wie das Root-Zeichen, die eine recht gute Differenzierung zulassen, die diagnostische Güte wird allerdings durch eine Kombination mehrerer Deskriptoren erhöht. Das Ziel, eine für sämtliche Läsionen zufriedenstellende Genauigkeit zu erreichen sowie eine gewisse Anzahl sicher bös- oder gutartiger Befunde zu detektieren, lässt sich ohne die sinnvolle Kombination der Ausprägungen verschiedener Deskriptoren schwerlich erreichen.

Man kann eine Dignitätsfestlegung aufgrund bestimmter Merkmalskombinationen auf verschiedenerelei Art und Weise realisieren. Empirische Scores haben hierbei den Vorteil einer einfachen Anwendung. Aufgrund der Prävalenz einer Eigenschaft innerhalb einer Dignität wird einem Merkmal ein Punktwert zugewiesen und letztlich eine Punktschme für die betrachteten Deskriptoren gebildet, aufgrund derer eine Malignitätsaussage getroffen wird.

Die Vergabe von Punktwerten gibt zwar gewisse Erfahrungswerte wieder, sie sind letzten Endes aber mehr oder weniger willkürlich und spiegeln nicht mathematisch exakt unterschiedliche Häufigkeiten einzelner Merkmalsausprägungen bei malignen und benignen Befunden wieder.

Eine weitere Möglichkeit ergibt sich durch Nutzung von Klassifikationsanalysen. Durch einen bestimmten Algorithmus wird hierbei ein Diagramm erstellt, welches einen Entscheidungsbaum darstellt, der zu einer bestimmten Diagnose führt. Die Anwendung ist für den Untersucher sehr praktikabel: Durch sequentielle Betrachtung bestimmter Deskriptoren und der Auswahl von deren Ausprägung wird der Befunder zu einer Diagnose geleitet. Am Endpunkt des Algorithmus ergibt sich eine bestimmte Merkmalskombination, aus der sich auch die Wahrscheinlichkeit für das Vorliegen einer gut- oder bösartigen Läsion ablesen lässt (Breiman 1984).

Durch die in dieser Arbeit angewandte binäre logistische Regression wird ebenfalls eine Wahrscheinlichkeit für das Vorliegen einer bestimmten Dignität errechnet. Vorteil beider Verfahren ist, dass man aufgrund der Angabe eines Wahrscheinlichkeitswertes exakt ablesen kann, wie sicher die Dignitätsvorhersage ist, wohingegen es sich bei empirischen Scores um ordinale Punktwerte handelt. Desto höher der Score ist, desto höher ist das Risiko für das Vorliegen eines Malignoms, mathematisch exakt kann man aber nicht sagen, um wie viel wahrscheinlicher ein Malignom pro Punktwert ist.

Die Anwendung der binären logistischen Regression erscheint auf den ersten Blick weniger praktikabel, für jede Deskriptorausprägung existiert eine spezielle Dezimalzahl, die in eine Gleichung zur Errechnung der Malignomwahrscheinlichkeit eingesetzt werden muss (siehe Kapitel 9.1.1). Zum einen ist dieses Verfahren umständlich, andererseits können sich durch Rechenfehler falsche Wahrscheinlichkeiten ergeben. Durch Implementierung der Rechenprozedur in ein geeignetes Softwaretool, in welchem man die Ausprägung für die Deskriptoren auswählen kann und daraufhin die Wahrscheinlichkeit berechnet wird, lässt sich dieser Nachteil beseitigen.

Bei adäquater Nutzung unterscheiden sich die drei Verfahren in der Schwierigkeit der Anwendung nicht wesentlich voneinander.

6.2 Methodische Aspekte

Im Rahmen dieser Arbeit kam es zum Einschluss von 1012 Patientinnen mit 1084 histologisch verifizierten Läsionen. Im Hinblick auf die Läsionsanzahl ist dies nach unserem Wissen das bisher umfangreichste Patientengut, mit dem ein Interpretationsmodell für die Bewertung von MR-Befunden erstellt wurde.

Die Indikationsstellung für die MR-mammographische Untersuchung erfolgte gemäß aktuell anerkannter Richtlinien (Mann et al. 2008). Mehrere Publikationen beschäftigen sich mit dem Einfluss von interventionellen Maßnahmen auf die Mammabildgebung (Fischer et al. 1996, Rieber et al. 1997, Morakkabati et al. 2003). Eine Beeinflussung der diagnostischen Qualität wird dabei kontrovers diskutiert. Da sich die vorliegende Arbeit speziell mit dem Einfluss morphologischer und KM-dynamischer Eigenschaften suspekter Anreicherungen befasst, wurden zahlreiche Befunde ausgeschlossen, um die Abhängigkeit der Zeichenbewertung von diversen Artefakten zu minimieren.

So wurden MR-Untersuchungen, denen Manipulationen wie eine OP, Stanzbiopsie oder Bestrahlung der Mamma sowie Chemotherapie weniger als sechs Monate vor Untersuchung vorausgingen nicht in die Studie eingeschlossen, da auf diese Art und Weise eine mögliche iatrogene Beeinflussung der morphologischen und kinetischen Läsionseigenschaften vermieden werden sollte. Diese Verfahrensweise erschien angemessen, da man mit dieser Arbeit zahlreiche Zeichen einer Bewertung zuführte, die bisher nicht ausreichend bzw. selten evaluiert wurden und keine sichere Aussage getroffen werden kann, inwiefern die angesprochenen diagnostischen und therapeutischen Maßnahmen die MR-mammographische Bildgebung beeinflussen könnten.

Aus der divergierenden Anzahl untersuchter Patientinnen und beschriebener MR-Befunde wird ersichtlich, dass auch mehr als ein Befund je Patientin Eingang in die Auswertung fand. Aufgrund ähnlicher Tumorbiologien bei einer Probandin kann es daher zu statistischen Verschiebungen kommen. Überschätzt werden sollte dies jedoch nicht, da von weniger als 10% der Untersuchten mehr als eine histologisch verifizierte Läsion in die Studie eingeschlossen wurde.

Zahlreiche verschiedene Untersuchungsparameter wie Art und Menge des injizierten KM, Feldstärke, Schichtdicke, Auflösungsvermögen, Flip-Winkel, Repetitionszeit, Echozeit,

2D- oder 3D-Gradientenechosequenzen, dargestellte Sequenzen und deren zeitlicher Ablauf zueinander wie auch viele weitere Faktoren offenbaren eine große Variationsbreite in der Bildgebung der MRM.

Die Kontrastmitteldosis belief sich bei allen Untersuchungen auf 0,1 mmol Gd-DTPA intravenös (Magnevist®, Schering Deutschland GmbH, Berlin, Deutschland) pro kg Körpergewicht, was dem Standard vieler Studien auf dem Gebiet der MR-Mammographie entspricht und auch aktuellen Empfehlungen gerecht wird (Knopp et al. 2003, Kuhl 2007a). Es existieren auch Arbeiten, bei denen doppelte KM-Mengen appliziert wurden (Heywang-Köbrunner et al. 1994). Bedingt durch die stärkere Anreicherung des Parenchyms soll jedoch die Differenzierung zwischen Gut- und Bösartigkeit erschwert sein, dies würde zu Lasten der Spezifität gehen (Kuhl 2007a).

Ausgehend von der Hypothese, dass eine gute diagnostische Aussagekraft mit der kombinierten Beurteilung morphologischer und kinetischer Parameter erzielt werden kann, erscheint es sinnvoll, eine jeweils akzeptable räumliche und zeitliche Auflösung anzustreben. Hierbei gilt der Grundsatz, dass eine verbesserte räumliche Auflösung zu Lasten der zeitlichen geht und umgekehrt (Kuhl 2007b).

Die Untersuchungen an unserem Institut erfolgten an vier verschiedenen Tomographen mit einer magnetischen Feldstärke von 1,5 Tesla. Der Untersuchungsablauf war durch ein standardisiertes Messprotokoll vorgegeben. Derzeit ist auch die Bildgebung mit 3-Tesla-Tomographen in der Erprobung, wobei man davon ausgeht, dass eine bessere Erfassung morphologischer Details von Raumforderungen möglich sein könnte (Kuhl et al. 2006).

Die Schichtdicken der von uns verwendeten MR-Geräte variierten geringfügig: Bei dem Philips Gyroscan, Philips Intera und dem Siemens Vision betrug sie 4 mm, beim Siemens Symphony 3 mm. Wegen des Distanzfaktors von 10% kamen jeweils 0,4 bzw. 0,3 mm zwischen den Schichten nicht zur Darstellung. Dies ist jedoch nicht relevant, wenn man bedenkt, dass Karzinome aufgrund ihrer Tumorangiogenese theoretisch erst ab 3 mm zur Darstellung kommen (Kaiser 1989). Außerdem berichten Warren und Mitarbeiter von einer Irrelevanz der verwendeten Schichtdicke hinsichtlich der diagnostischen Genauigkeit der MRM (Warren et al. 2009).

Einer gewissen Schwankungsbreite in der T1- und T2-Wichtung unterlag geräteabhängig auch die Bildmatrix und somit die Bildauflösung. Der Nutzen räumlich hochauflösender MRTs wird kontrovers diskutiert, von Obenauer et al. wird er als gering beschrieben (Obenauer et al. 2002). In logischer Konsequenz müsste dann auch eine höhere diagnostische

Güte von 3-Tesla-Tomographen nachgewiesen werden, da mit diesen eine bessere Bildgebung möglich erscheint.

Bezüglich der zeitlichen Abfolge der Messsequenzen erfolgte bei uns die Anfertigung der T2-Aufnahmen nach der KM-Applikation. So konnten die besonders bewegungsempfindlichen dynamischen Messungen früher erfolgen, was insbesondere für die Erfassung quantitativer Daten im Rahmen weiterer Forschungsarbeiten des IDIR als notwendig erachtet wurde. Vom Einfluß dieser Vorgehensweise wird noch in Kapitel 6.3 bei der Erörterung des Deskriptors *T2-Helligkeit* die Rede sein.

Die Länge des in dieser Arbeit betrachteten Zeitraumes bedingt einen potentiellen Einfluß von technischen Weiterentwicklungen der MRT. So ermöglichen moderne Geräte eine höhere räumliche Auflösung und somit Detailerkennbarkeit im Vergleich zu älteren Tomographen. Die in dieser Studie verwendeten Untersuchungssequenzen wurden technisch sorgfältig so abgestimmt, dass Bildeindruck und Kontrast über die Jahre konstant blieben. Die in dieser Arbeit verwendeten Deskriptoren ließen sich somit einheitlich an allen Aufnahmen evaluieren. Eine möglicherweise verbesserte Erfassung morphologischer Charakteristika in kleinen Herdläsionen wurde in dieser Arbeit nicht betrachtet. Allerdings konnte unsere Arbeitsgruppe die Applizierbarkeit der hier beschriebenen Deskriptoren im verwendeten Untersuchungskollektiv an kleinsten Herdbefunden (≤ 5 mm) demonstrieren (Dietzel et al. 2009).

In die Auswertung gingen nur histologisch verifizierte Befunde ein. Anreicherungen, die mit höchster Wahrscheinlichkeit auf das Vorliegen einer benignen Veränderung hindeuteten, wurden nicht in jedem Fall histologisch abgeklärt. Auch vorbeschriebene Befunde, die über einen längeren zeitlichen Verlauf keine Veränderungen wie z. B. Größenzunahme aufwiesen, kamen nicht zwingend zu einer histologischen Abklärung. Weiterhin wurden Mammæ ohne MR-mammographische Anreicherungen nicht für die Modellerstellung berücksichtigt. In der Konsequenz ist somit eine Überrepräsentation maligner Befunde naheliegend, was durch ein Follow-up negativer Befunde über einen gewissen Zeitraum hätte vermieden werden können. Andererseits war eine größtmögliche sichere Zuweisung der Dignität zu einem MR-mammographischen Korrelat am ehesten durch die histologische Sicherung zu gewährleisten.

Bei Bewertung unterschiedlicher Interpretationsmodelle kann mit statistischen Kennzahlen wie u. a. Sensitivität, Spezifität, positiver bzw. negativer prädiktiver Wert, diagnostische

Genauigkeit und AUC die Qualität solcher Modelle verglichen werden. Besonders Sensitivität und Spezifität sind hierbei Werte, welche eine Aussage über die Güte eines Screening-Verfahrens liefern. Wenngleich man oft insbesondere Werte für Sensitivität und Spezifität im Zusammenhang mit publizierten MRM-Studien anführt, so muss auch dem Umstand Rechnung getragen werden, dass die meisten Studien keine Screeningkollektive rekrutiert hatten. Meist handelte es sich eher um Patientinnen mit suspekten Vorbefunden, die einer additiven diagnostischen Abklärung zugeführt wurden. Somit basiert die Mehrzahl der Daten auf Erkenntnissen, die aus Risikokollektiven gewonnen wurden.

Der PPV gibt die Wahrscheinlichkeit an, mit der ein positiv klassifizierter Befund auch tatsächlich positiv ist. Er ist von der Prävalenz der positiven Fälle in der Grundgesamtheit abhängig und umso höher, desto höher die Prävalenz ist (Altman und Bland 1994). Ebenso ist der NPV vom zugrunde liegenden Kollektiv abhängig. Will man also die Güte verschiedener Modelle hinsichtlich der prädiktiven Werte vergleichen, darf dies nicht ohne genauere Betrachtung der Zusammensetzung des Studienkollektivs geschehen. Im Gegensatz zu dieser Arbeit schließen diverse Autoren Fälle ein, bei denen in den MR-Aufnahmen kein anreichernder Prozess auffiel. Da auf der Basis pathophysiologischer Überlegungen fast jeder Tumor ab einer Größe von 3 mm KM anreichert, ist bei einer negativen Bildgebung mit einer an 100% grenzenden Wahrscheinlichkeit ein Malignom auszuschließen. Hieraus kann man schlussfolgern, dass sich bei Modellen, die auch MR-Untersuchungen ohne Befund einschließen, bessere statistische Kennzahlen ergeben. Dies wird u. a. die diagnostische Genauigkeit betreffen, aber auch die oft im Zusammenhang mit der MRM kritisch betrachtete Spezifität, da sich ein fehlender Befund wohl schlecht als falsch positiv zuordnen lassen würde.

Anhand bloßer Zahlen allein lassen sich dadurch unterschiedliche Studien nicht mit der vorliegenden Arbeit vergleichen. Behelfen kann man sich jedoch, indem man die im angesprochenen Fall eingeschlossenen Untersuchungen ohne Anreicherung ausschließt und dann erneut die statistischen Kennzahlen errechnet, wenn dies mit Hilfe der in den jeweiligen Studien angegebenen Ergebnisse möglich ist. So errechnen sich dann die aufgrund des Modells richtig-positiv und richtig-negativ klassifizierten Anreicherungen dividiert durch alle Befunde als neue diagnostische Genauigkeit, die dann mit den berechneten Modellen der vorliegenden Arbeit verglichen werden kann.

6.3 Diskussion der für die Regressionsmodelle verwendeten Deskriptoren

Die Deskriptorenanzahl zur Berechnung der Wahrscheinlichkeit eines Malignoms kann nach den Ergebnissen dieser Studie im Vergleich zu Deskriptormodellen unter Einschluss aller verfügbaren Zeichen ohne relevante Verminderung der diagnostischen Genauigkeit deutlich reduziert werden. Signifikante Unterschiede zwischen den zugehörigen AUC sind durch hohe Fallzahlen bei kleinen Unterschieden bedingt. Letztendlich ist abzuwägen zwischen Einfachheit des Interpretationsmodells mit lediglich fünf Deskriptoren und einem komplexeren Modell, dessen diagnostische Genauigkeit etwas verbessert ist und mehr eindeutige Befunde liefern kann bei nicht wesentlich erhöhter Fehlerquote. Nachfolgend soll insbesondere die Aussagekraft der fünf modellabhängigen Deskriptoren bewertet werden, da nach den erhobenen Daten jeweils ein Modell mit reduziertem Deskriptorensatz in der Lage war, benigne oder maligne Befunde mit hoher Sicherheit zu klassifizieren.

Gemeinsam ist den Deskriptoren die rein visuell-qualitative Beurteilung. In diesem Zusammenhang sollte stets die Reliabilität der erhobenen Daten gesichert sein. Quantitative computergestützte Methoden wurden nicht verwendet. Allerdings bleiben gängige Publikationen den Nachweis schuldig, dass damit eine verbesserte Diagnostik ermöglicht wird. Verbreitet ist die (semi-) quantitative Bewertung des Signalintensitätskurvenverlaufes. Der Untersucher setzt dabei mehrere ROIs (region of interest) in die am stärksten enhancenden Bezirke der Läsion, woraus computerbasiert die Höhe der Signalintensität durch eine Kurve über den zeitlichen Verlauf dargestellt wird. Ebenso existieren Computer-aided-diagnosis (CAD)-Systeme, welche die Kinetik unterschiedlicher Bezirke innerhalb einer Raumforderung darstellen und quantifizieren können (Baltzer et al. 2009a).

Solche Verfahren sind zwar von subjektiven Einflüssen befreit, eine verbesserte Diagnostik wird jedoch angezweifelt (Kinkel et al. 2000, Baltzer et al. 2009a). Die Erfassung quantitativer Messwerte bedingt in der Konsequenz nicht zwingend reproduzierbare Werte. Eine Studie ermittelte quantitative Parameter an unterschiedlichen Tomographen bei identischen Messparametereinstellungen. Trotzdem ergaben sich signifikante Abweichungen der gemessenen Werte (Pabst et al. 2001). Nachfolgend werden bei einigen Zeichen zwecks Nachvollziehbarkeit der Interobserverreliabilität (IOR) in der gängigen Literatur angegebene Kappa-Werte aufgeführt. Für die Interpretation von Kappa-Werten (κ) ist folgender Vorschlag geläufig: Ein $\kappa < 0$ spräche für „schlechte Übereinstimmung (poor agreement)“, κ zwischen 0 und 0,20 für „etwas (slight) Übereinstimmung“, 0,21 bis 0,40 für „ausreichende (fair)

Übereinstimmung“, 0,41 bis 0,60 für „mittelmäßige (moderate) Übereinstimmung“, 0,61-0,80 für „beachtliche (substantial)“ und 0,81 bis 1,00 für „(fast) vollkommene ([almost] perfect) Übereinstimmung“ (Landis und Koch 1977).

In der vorliegenden Arbeit wurde aufgrund des Langzeit-Studienkonzepts mit Konsensusbewertung durch zwei erfahrene Radiologen die Reliabilität der einzelnen MRT-Deskriptoren nicht adressiert. Die IOR in dieser Untersuchung wurde durch einheitliche Schulung der Untersucher an der gleichen Institution sowie eine der Datenerfassung vorangehende Trainingssession minimiert.

Zu beachten ist bei dieser Arbeit, dass die Höhe der initialen Signalintensität in Form eines schwachen, mittleren oder starken Signalanstieges und deren diagnostische Bedeutung nicht evaluiert wurde.

Bei dem Modell, welches Vorteile bei der Selektion sicher maligner Veränderungen besitzt, wurden die Deskriptoren *zeitliches* und *räumliches KM-Anreicherungsverhalten*, *Root-Zeichen*, *Ödem* und *T2-Helligkeit* genutzt. Dabei handelt es sich sowohl um herdassoziiert morphologische und kinetische Kriterien sowie mit der Beurteilung auf Vorliegen ödematöser Auffälligkeiten um einen Parameter, der eine Veränderung des den suspekten Prozess umgebenden Mammaparenchyms erfasst. Insgesamt erfolgte somit die Auswahl aus einem Deskriptorvariablenpool aller an 1084 Läsionen evaluierten Zeichen. Das Modell, welches die Ausprägung von *zeitliches* und *räumliches KM-Anreicherungsverhalten*, *Berandung*, *Form* und *Septen* berücksichtigt, wurde aus durch mehrere Arbeiten etablierten Deskriptoren zusammengestellt, um einen Modellvergleich zu dem zuvor Errechneten zu ermöglichen. Die Selektion der Parameter für das optimierte Fünf-Deskriptoren-Modell geschah entsprechend der Vorwärtsselektion der binären logistischen Regression. Einige Deskriptoren wurden dabei in der Reihenfolge der Variablenaufnahme hintenangestellt, obwohl sie gute statistische Kennzahlen wie z. B. den der diagnostischen Genauigkeit aufwiesen. Es ist davon auszugehen, dass sich aus der Aussagekraft bestimmter Deskriptoren ein mehr oder weniger ähnlicher Informationsgehalt ergibt. So ist denkbar, dass Deskriptoren wie *Rand*, *Blooming* und *Root-sign*, welche den Läsionsrand charakterisieren, bei gemeinsamer Berücksichtigung ein Modell nicht wesentlich verbessern und die Erfassung einer dieser Deskriptorvariablen für die Modellerstellung ausreichend erscheint.

Die vorliegende Arbeit befasste sich mit herdförmigen Anreicherungen, welche potentiell einen malignen Befund präsentieren konnten. Eine Differenzierung zwischen sich räumlich abgrenzenden raumfordernden Herdbefunden und infiltrativen die anatomischen Grenzen akzeptierenden Prozessen (auch non-mass-Läsionen oder Infiltrationen), so wie sie das BI-RADS-Lexikon (ACR 2003) vorschlägt, erfolgte nicht (siehe Abb. 6.1). Neuere Veröffentlichungen folgen zunehmend besagter Differenzierung. Diese Studie hatte zum Ziel, eine erreichbare diagnostische Genauigkeit mithilfe der am IDIR über einen größeren Zeitraum untersuchten Parameter zu ermitteln und herauszufinden, wie häufig eine diagnostisch sichere Aussage möglich ist. Einige in dieser Arbeit beschriebene Zeichen sind typischerweise mit soliden Herdbefunden assoziiert und implizieren somit die Eigenschaft eines Herdes. Dazu gehören Root- und Hook-Zeichen, ebenso sind eine lobulierte Form oder das Auftreten von Septen bei nicht soliden Anreicherungen undenkbar. Eine Differenzierung zwischen soliden und nicht soliden Anreicherungen wurde in dieser Arbeit aufgrund des langen Studienzeitraums nicht vorgenommen, da zu dessen Beginn die BI-RADS-Klassifikation noch nicht vorgeschlagen worden war. Bisher wurde eine diagnostische Überlegenheit einer Differenzierung zwischen soliden und nicht soliden Herdbefunden nicht demonstriert, obwohl die Einteilung nach radiologischen und biologischen Kriterien deskriptiven Sinn macht. Eine unterschiedliche Aussagekraft von gleichen Deskriptoren in beiden morphologischen Klassen ist aufgrund obiger Ausführungen durchaus denkbar.

Für die Kontrastmitteldynamik nach initialem Signalanstieg bei nicht herdbefundartigen Anreicherungen konnte bisher kein Nutzen nachgewiesen werden, sowohl was die zeitliche Konstanz der KM-Anreicherung als auch die räumliche im Sinne einer homogenen bzw. heterogenen angeht, wohingegen die diagnostische Bedeutung bei Raumforderungen unumstritten ist (Jansen et al. 2008, Yabuuchi et al. 2009). Entsprechend könnte die Debatte darauf ausgedehnt werden, ob ein perifokales Ödem oder zuführendes Gefäß für ein Malignom bei Herdbefunden, bei non-mass-Läsionen oder bei beiden Spezifikationen spräche oder nicht.

Es muss auch die Frage gestellt werden, inwieweit mehrere Befunder eine sichere Zuordnung zum Läsionstypen treffen können. Eine Publikation von Ikeda et al. beschreibt einen κ -Wert von 0,57 für die Zuordnung des Läsionstyps, was eine mittelmäßige Übereinstimmung anzeigt (Ikeda et al. 2001).

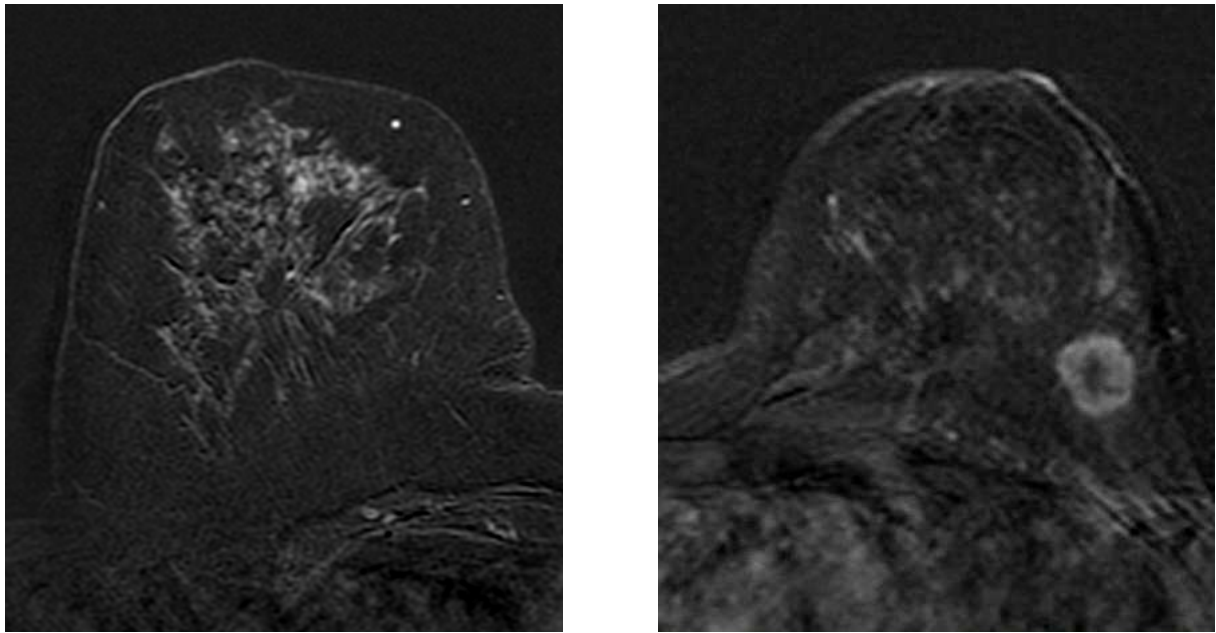


Abbildung 6.1: Non-mass-Läsion und Herdbefund

Links Darstellung einer sich diffus ausbreitenden non-mass-Läsion, rechts ein typischer Herdbefund. Beide Aufnahmen wurden 7 min nach KM-Injektion angefertigt.

Zeitliches KM-Anreicherungsmuster

Wie in vielen anderen Studien war die Betrachtung der KM-Anreicherung im zeitlichen Verlauf nach initialem Signalanstieg auch nach den Ergebnissen der vorliegenden Arbeit ein wichtiges diagnostisches Werkzeug (Kaiser und Zeitler 1989, Kaiser 1993, Boetes et al. 1994, Kuhl et al. 1999b). Eine persistierende Zunahme der Signalintensität war ein guter Indikator für eine benigne Läsion, ein wash-out sprach für einen bösartigen Prozess. Bei Vorliegen eines Plateau-Phänomens ließ sich hingegen keine sichere Aussage treffen, je ein Viertel der gutartigen bzw. bösartigen Histologien zeigten eine Plateau-Kinetik.

Problematisch bei Beurteilung der zeitlichen KM-Dynamik hätte die visuelle Auswertung sein können, wenn geringe Signalunterschiede für den Befunder nicht wahrnehmbar gewesen wären. Kinkel et al. beschreiben einen κ -Wert von 0,80 für das Vermögen von Bildbefundern, zwischen dem Vorliegen oder der Abwesenheit eines wash-out zu unterscheiden und einen κ -Wert von 0,69 für die generelle Charakterisierung der zeitlichen Kontinuität des Signalintensitätsverhaltens, wobei sich diese Arbeitsgruppe allein auf die visuelle Auswertung am Ort des maximalen Signalanstiegs konzentrierte. Auf das Festlegen von ROIs und die

daran anschließende computerbasierte Auswertung wurde somit verzichtet (Kinkel et al. 2000).

Es stellt sich die Frage, ob eine weniger subjektive Evaluierung der Kontrastmittelkinetik vorteilhaft für die Befundung sein könnte. Hierzu existiert der Vorschlag, die am stärksten enhancenden Bereiche als ROIs zu kennzeichnen und durch eine Software eine Signalintensitätskurve entwerfen zu lassen, aus der man den zeitlichen Verlauf der Signalstärke ablesen oder berechnen kann (Kuhl et al. 1999b). Die Untersuchung der Kuhl-Arbeitsgruppe fand eine sehr gute Übereinstimmung bei der visuellen Bewertung von Signalintensitäts-Zeitkurven. Jedoch soll die manuelle untersucherabhängige Festlegung von ROIs einer zu starken Variabilität unterliegen (Mussurakis et al. 1996, Kinkel et al. 2000). Insgesamt scheinen visuelle Beurteilung, Signalintensitätskurvenverlaufsbeurteilung und Analyse durch computerbasierte Methoden diagnostisch gleichwertig zu sein (Baltzer et al. 2009c).

Eine andere Arbeit untersuchte die Notwendigkeit der Betrachtung des zeitlichen Verlaufs der Signalintensität für die gesamte Läsion. Die Auswertung erfolgte unter prozentualer Angabe der Bereiche der Raumforderung, welche KM weiter anreicherten, es wieder auswuschen oder in der KM-Aufnahme sistierten. Es handelt sich hierbei um ein sog. CAD-System. Ein zusätzlicher Nutzen für diese Vorgehensweise konnte jedoch nicht nachgewiesen werden, so dass unsere Methodik für die Bewertung des hier diskutierten Merkmals als völlig ausreichend betrachtet werden kann. Selbige Arbeit bezieht auch die initiale KM-Aufnahme mit ein, eine signifikante Verbesserung hinsichtlich der diagnostischen Genauigkeit ergab sich jedoch nicht. Insgesamt genügt es also, den initial am stärksten anreichernden Teil einer Läsion zur Evaluation des hier diskutierten Merkmals zu betrachten (Baltzer et al. 2009a).

Weiterhin bleibt fraglich, ob die Analyse kinetischer Daten für Herdbefunde und non-mass-Läsionen gleichermaßen effektiv ist. Eine Studie konnte keinen Nutzen bei der Auswertung kinetischer Informationen von nicht herdbefundartigen Anreicherungen feststellen (Jansen et al. 2008).

Räumliches KM-Anreicherungsmuster

Das räumliche KM-Anreicherungsmuster wurde im Rahmen der vorliegenden Arbeit in die Kategorien *homogen*, *heterogen*, *zentrifugal* und *zentripetal* untergliedert. Für die zentripetale KM-Dynamik werden in vielen anderen Arbeiten Begriffe wie *Ringenhancement*, *Rim-sign*, *peripher*, *peripheres Ringenhancement* oder gar *zirkumferential* synonym verwendet.

Im Einklang zu zahlreichen anderen Quellen war die zentripetale Anreicherungscharakteristik (vgl. Abb. 6.2) mit einer Häufigkeit von einem Drittel bei den Malignomen mit einem PPV von 87,8% ein sehr starker Prädiktor für Bösartigkeit (Schnall et al. 2006, Tozaki et al. 2006, Yabuuchi et al. 2008). Histologisch wird das Phänomen durch eine größere Gefäßdichte in der Tumorperipherie als im Tumorzentrum erklärt, in dem stattdessen regressive Veränderungen wie Fibrosen oder Nekrosen vorherrschen und eine Anreicherung mit zeitlicher Verzögerung auftritt (Buadu et al. 1996).

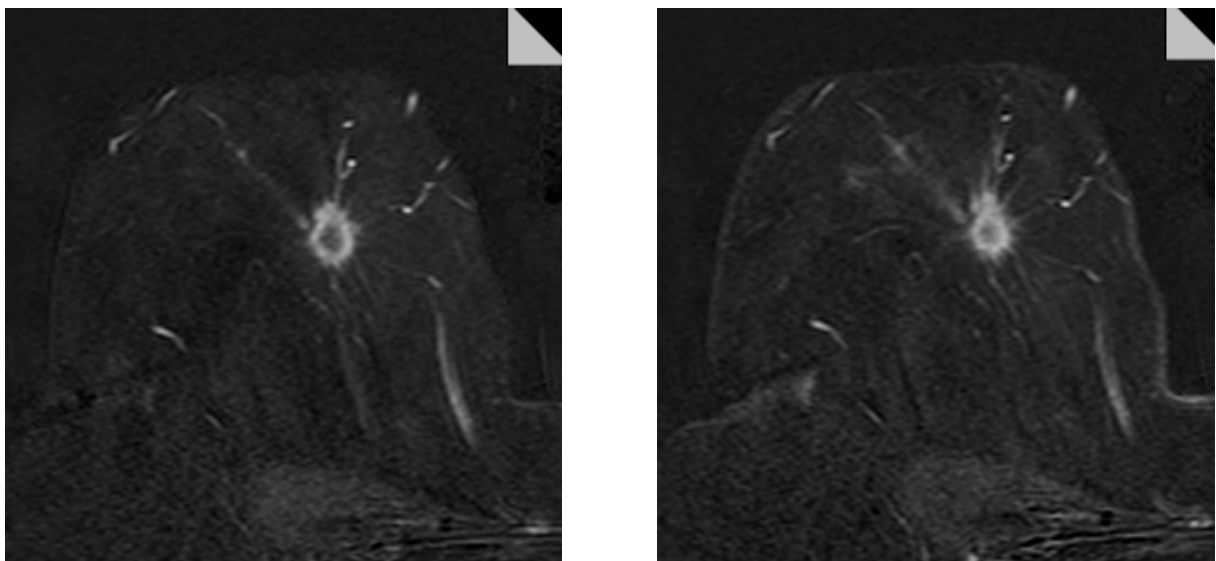


Abbildung 6.2: Zentripetales Enhancement

T1-Subtraktionsaufnahmen 1 min (li.) und 7 min (re.) nach KM-Applikation mit von der Peripherie ins Zentrum fortschreitendem KM-Anreicherungsmuster

54% der bösartigen Tumoren präsentierten sich mit einem *heterogenen Anreicherungsmuster*. Trotzdem kann man bei Zutreffen dieses Charakteristikums nicht auf die Dignität schließen, da sich 55% der gutartigen Veränderungen ebenfalls so darstellten.

Die *zentrifugale Anreicherung* ist nur sehr spärlich Gegenstand von Publikationen, die sich mit der MRM auseinandersetzen. Auch im BI-RADS-Katalog des ACR taucht diese Merkmalsbeschreibung nicht auf. Stattdessen findet dort ein *zentrales Enhancement* Erwähnung, das mit einer zentrifugalen Anreicherung nicht verwechselt werden darf. Es soll sich um ein „Nest“ innerhalb eines Herdes handeln und gewöhnlich stärker als der Rest der Raumforderung anreichern. Meist sollen ein höhergradiges duktales Karzinom oder vaskuläre Brusttumoren ein solches Anreicherungsmuster aufweisen (Erguvan-Dogan et al. 2006).

Eine zentrifugal anreichernde Struktur ist hingegen eher als gutartig zu bewerten (Boetes et al. 1994). 1% der bösartigen und 3% der gutartigen Befunde zeigten eine entsprechende Anreicherungsdynamik, was auch zeigt, dass dieses Zeichen sehr selten auftritt.

Ein wichtiger Hinweis auf Gutartigkeit ist das homogene KM-Anreicherungsverhalten, welches ca. ein Drittel der gutartigen Befunde und 13% der Malignome aufwiesen.

Alles in allem stellt das räumliche KM-Anreicherungsmuster ein gutes und durch mehrere Studien etabliertes diagnostisches Element in der Dignitätsbeurteilung dar. Zu kritisieren ist die wohl schlechte Interobserverreliabilität bei einem κ -Wert von 0,22 (Kinkel 2000), wobei für die mögliche Ausprägung des Binnenanreicherungsmusters Kriterien verwendet wurden, die von denen der vorliegenden Arbeit abweichen.

Rand

Kritikpunkt dieser Arbeit könnte die vereinfachte Klassifizierung des Randes sein. Das BI-RADS-Manual empfiehlt eine Differenzierung zwischen gut definierter sowie schlecht definierter Berandung und fordert überdies eine Bewertung, ob radiäre Ausziehungen (*spikuliert*) existieren. Damit kommt man auch dem üblichen Procedere in der Röntgenmammographie nach (Lieberman et al. 1998).

Eine Vielzahl von Publikationen hat sich bereits mit der Berandung und der daraus resultierenden Schlussfolgerung hinsichtlich der Dignität einer Läsion beschäftigt. Ein spikulierter Rand ist stets als hochsuspekt zu werten, gute Abgrenzbarkeit zur Umgebung spricht hingegen eher für Benignität (Nunes et al. 1997b, Nunes et al. 1997a).

Die Bedeutung einer spikulierten Berandung lässt sich durch die für dieses Zeichen publizierten positiven prädiktiven Werte ermessen, die studienabhängig zwischen 76 und

100% liegen (Nunes et al. 1997b, Liberman et al. 2002, Tozaki et al. 2006). Die Beurteilung des Läsionsrandes ist als wichtiges diagnostisches Hilfsmittel einzuschätzen, nicht zuletzt wird auch von einer zufriedenstellenden Reliabilität bei der Befundung berichtet (Kinkel et al. 2000, Ikeda et al. 2001). Es existieren kaum Interpretationsmodelle oder Scoring-Systeme, aus denen die Randbewertung wegzudenken wäre (Nunes et al. 1997a, Kinkel et al. 2000, Baum et al. 2002, Szabo et al. 2003, Schnall et al. 2006). Im Interesse einer uniformen MR-Befundung sollte künftig eine Einteilung nach scharfen, unscharf begrenzten und spikulierten Läsionen vorgenommen werden. Es bleibt eine offene Frage, ob der Deskriptor *Rand* eine wichtigere Rolle bei den in dieser Arbeit betrachteten Interpretationsmodellen gespielt hätte, wenn man obige Differenzierung in drei Kategorien vorgenommen hätte.

Ödem

Die Detektion von Flüssigkeit zum Nachweis tumorassoziierter Ödeme in T2-gewichteten Aufnahmen macht man sich bereits bei der Diagnostik von Hirn- und Weichteilgewebstumoren zunutze (Holland et al. 1985, Kalayanarooj 2008). Die Betrachtung von T2w-Aufnahmen hinsichtlich des Vorhandenseins ödematöser Veränderungen scheint ebenfalls ein vielversprechendes diagnostisches Werkzeug für die MR-Mammographie zu sein.

Eine Kausalität zwischen Ödementstehung und Tumor erklärt man sich durch das invasive Tumorstadium, welches mit gesteigerter tumorperipherer Proteolyse und vermehrter Neoangiogenese in Verbindung mit erhöhter Kapillarpermeabilität durch basalmembranäre Lecks einhergeht. Zusätzliche Erklärungsansätze beschreiben eine durch Entzündungsmediatoren induzierte kapillare Permeabilitätserhöhung bei begleitenden inflammatorischen Prozessen, aber auch lymphatische Abflussstörungen werden diskutiert (Dvorak et al. 1981, Senger et al. 1990, Bässler 1997, Yang 2006).

Abgesehen von einer genaueren Untersuchung an unserer Einrichtung durch F. Yang, dem Vorschlag einer Evaluierung durch Malich et al. und einer Untersuchung inflammatorischer Mammakarzinome ist von der Bedeutung, die ein läsionsassoziertes Ödem für die Dignitätsaussage haben könnte in der gängigen Literatur kaum die Rede (Malich et al. 2005, Yang 2006, Renz et al. 2008). Im Fokus steht insbesondere die perifokale Lokalisation, da die Ödem-hervorrufenden Prozesse speziell an den Randzonen des Tumors wirken sollen (Bässler 1997). Weiterhin können um Tumorausläufer befindliche Flüssigkeitseinlagerungen als diffus-ipsilaterale Signalanstiege im T2-Bild imponieren.

Grundsätzlich kann eine Läsion mit begleitendem Ödem, vor allem wenn dieses ein perifokales ist, als malignomsuspekt gewertet werden. Bei der Interpretation von Ödemen ist allerdings die Patientengeschichte in Bezug auf vergangene diagnostische und therapeutische Eingriffe zu beachten. Bei Zustand nach Radiatio sind ödematöse Veränderungen ein typischer Befund (Morakkabati et al. 2003).

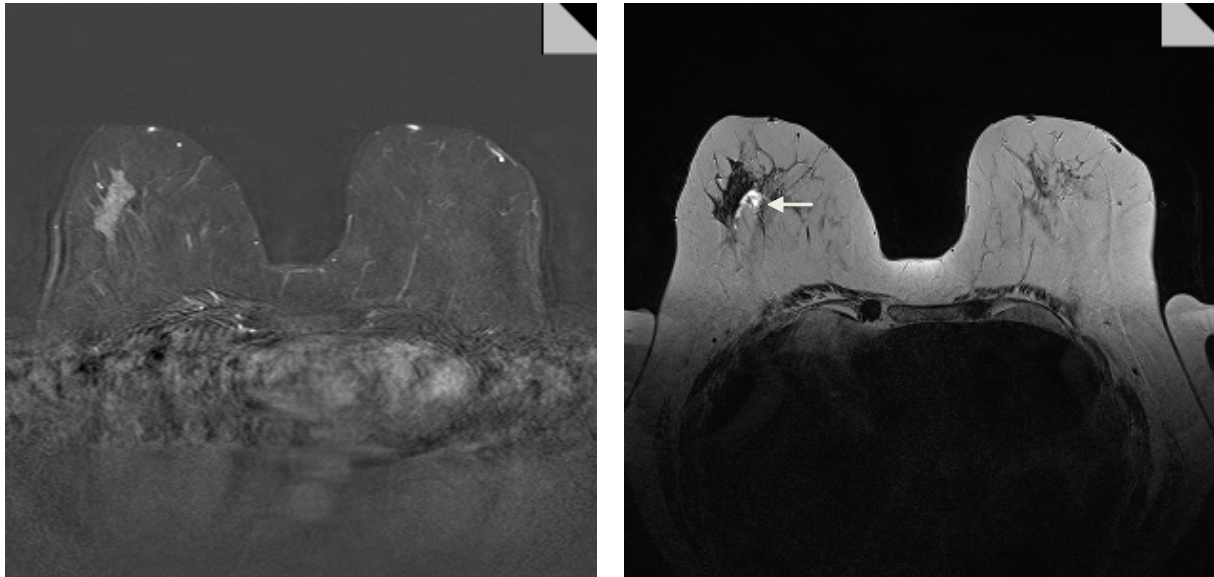


Abbildung 6.3: Perifokales Ödem

Links Subtraktionsaufnahme, rechts T2w-Bild mit perifokal verlängerter T2-Zeit entsprechend einer pathologischen Flüssigkeitsvermehrung im Gewebe

T2-Helligkeit

Die Signalintensität in den T2-gewichteten Aufnahmen ergibt sich vorrangig durch Fibrosierungsgrad sowie Wasser- und Fettgehalt des betrachteten Gewebebezirkes. Hyperintens stellt sich dementsprechend der Formenkreis der zystischen Befunde dar, zu dem einfache Zysten, intrazystische Karzinome und intrazystische Papillome zuzuordnen sind, aber auch schleimhaltige Tumoren und flüssigkeitsreiche Veränderungen wie Abszesse, Ödeme und Nekrosen (Kawashima et al. 2002). Die erhöhte Signalintensität bei Schleimanteilen kommt durch die Glykoproteinmatrix zustande, welche vermehrt Wasser bindet. Aber auch eingeschmolzene Nekrosen von Tumoren bewirken zentral einen T2-Signalanstieg (Kuhl et al. 1999a, Kawashima et al. 2002).

Ergebnisse von Kuhl et al. zeigten, dass 87% der Malignome iso- oder hypointens und 71% der Fibroadenome hyperintens waren (Kuhl et al. 1999a). Unseren Daten zufolge imponierte

nur ein marginaler Anteil von 3% der bösartigen Läsionen hyperintens, mit 22% war der Anteil bei den gutartigen Pathologien wesentlich höher. Letztendlich lässt sich bei iso- oder hypointensen Befunden anhand der Beurteilung dieses Deskriptors allein keine sichere Diagnose ableiten. Bei Hyperintensität ist von einem gutartigen Prozess auszugehen, insbesondere dann, wenn auch andere MR-Merkmale wie glatte Berandung, runde, ovale oder lobulierte Form und ein Vorhandensein dunkler Septen darauf hindeuten (Hochman et al. 1997). Differentialdiagnostisch muss jedoch dem Umstand Beachtung gezollt werden, dass sich auch muzinöse Karzinome oder Tumornekrosen hyperintens darstellen (Bick 2000). Zusätzlich ist der zeitliche Zusammenhang zwischen KM-Applikation und Anfertigung der T2-Aufnahmen zu betrachten. Unsere T2-Sequenzen wurden ausnahmslos nach den dynamischen Serien gemessen, während andere Arbeitsgruppen wie um Kuhl die T2-Messung vor den T1-Scans durchführten (Kuhl et al. 1999a). Dieser Aspekt scheint nicht völlig unbedeutend, da Wielopolski et al. einen Signalabfall von bis zu 30% bei Anwendung eines dem unsrigen vergleichbaren Messprotokolls fanden (Wielopolski 2001). Auch in einer anderen Arbeit wurden Phantommessungen durchgeführt, die einen 10%igen Signalabfall in der T2w bei den zu erwartenden KM-Konzentrationen im Tumor erwarten ließen. Letztendlich ist der Signalabfall um so höher, desto stärker der Prozess anreichert (Baltzer et al. 2009b). Eine Vergleichbarkeit des an dieser Stelle diskutierten Deskriptors mit anderen Arbeiten ist somit erschwert. Trotzdem wird diesbezüglich der Zeitpunkt der T2-Messung an unserem Institut als weniger bedeutend erachtet, da die T2w-Aufnahmen vor Studienbeginn vor der KM-Dynamik angefertigt wurden und sich auch schon damals Karzinome erfahrungsgemäß hypointens darstellten.

Root- und Hook-Zeichen

Bei Root- und Hook-Zeichen bezieht man sich auf Tumorausläufer ins umgebende Gewebe unter der Annahme eines infiltrierenden Karzinomwachstums, beim Hook-sign speziell auf einen Ausläufer, der die präpektoriale Fettschicht durchbrochen hat.

Man kann diese Interpretation in Anlehnung an die Röntgenmammographie sehen, bei der ebenfalls Ausläufer ins umgebende Gewebe als karzinomverdächtig gewertet werden (Tardivon et al. 2002).

Bemerkenswert ist, dass sich das Root-sign, aber auch das Hook-sign in unserer Untersuchung als bedeutende diagnostische Parameter herausstellten. Beide sind leider bisher

wenig untersucht, und es lässt sich auch keine Aussage darüber treffen, inwieweit verschiedene Observer das Vorhandensein eines Root- oder Hook-signs übereinstimmend bewerten würden.

Das histopathologische Erscheinungsbild dieser beiden Zeichen lässt sich noch nicht eindeutig nachvollziehen. Meist versucht man die Erklärung durch eine desmoplastische Veränderung, also einer fibrösen Antwort des Nachbargewebes auf den tumorösen Prozess (Walker 2001).

Weitere Ursachen sind zu diskutieren. So kann man postulieren, dass als Cooper-Ligamente bezeichnete Retinakula, welche das Brustdrüsengewebe durchziehen und es mit der Fascia pectoralis verbinden, als desmoplastische Reaktionen fehlgedeutet werden könnten, da sich die MR-morphologischen Korrelate ähneln. Auch narbige Umbauprozesse, wie sie postoperativ, nach Entzündungen oder infolge mastopathischer Veränderungen auftreten, könnten verantwortlich sein.

Kaiser beschreibt das Hook-sign als feine fadenartige Verbindung zum Pektoralismuskel, die entweder wenig oder gar keine Kontrastmittelaufnahme zeigt (Kaiser 2007). Über die wahre Ätiologie des Hook-signs kann nur spekuliert werden. Die Qualität des Hook-Zeichens ergibt sich aus den von uns gefundenen statistischen Werten: Mit 96,8% und 93,1% wurden die höchsten Werte für Spezifität und PPV erreicht. Im Vergleich zeigt das Root-Zeichen eine deutlich höhere Prävalenz bei Malignomen als das Hook-Zeichen bei ähnlich guten statistischen Werten. In Abb. 6.4 wird nochmals auf den in Abb. 5.1 dargestellten Herdbefund Bezug genommen, in dem das Root-sign positiv war.

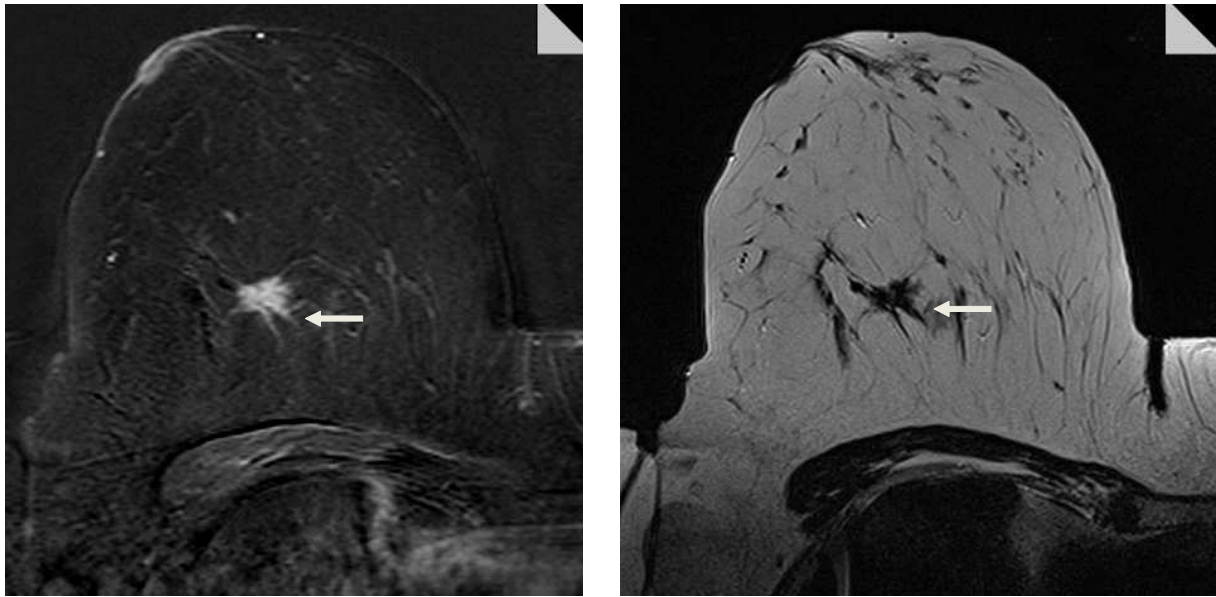


Abbildung 6.4: Root-sign

Links Subtraktionsaufnahme 7 min nach KM-Applikation, rechts T2w-Abbildung

Septen

Signalarme Septen sind ein allgemein anerkanntes Benignitätskriterium (Nunes et al. 2001, Tozaki et al. 2006). Diffizil stellt sich jedoch eine eindeutige Stellungnahme bei Vorhandensein anreichernder Septen dar. So beschreibt Tozaki einen PPV von 97% bei verspätet anreichernden läsionsinternen Septen, allein die Prävalenz solcher Strukturen liege bei den Malignomen bei ca. 75% (Tozaki et al. 2005). Anreichernden innerläsionären Septen sagt man nach, dass sie eher für Malignität sprechen (Erguvan-Dogan et al. 2006). Im Gegensatz zu Tozakis Arbeit wiesen bei uns lediglich 1% (8/648) der malignen Läsionen derartige Septierungen auf, bei den benignen Anreicherungen waren es mit 4% (16/436) sogar mehr.

Diverse andere Studien stimmen mit den von uns gefundenen Ergebnissen überein. U. a. untersuchten Kinkel et al. 34 Malignome, wobei eins anreichernde Septen aufwies (3% entsprechend). Bei den 23 gutartigen Prozessen fanden sich in drei Fällen (9% entsprechend) Septierungen, zwei reicherten an und das verbleibende Septum stellte sich dunkel dar (Kinkel et al. 2000).

Eine eindeutige Aussage über die Bedeutung anreichernder Septen lässt sich nur zurückhaltend treffen. Es stellt sich die Frage, ob diese Merkmalszuschreibung irreführend ist. Eine Arbeitsgruppe fand für das MR-mammographisch sichtbare Septenenenhancement bei der histologischen Aufarbeitung feine Tumorzellstränge (Qayyum et al. 2002). Eine genauere histologische Entsprechung dieses Merkmals ist notwendig.

Form

Die Formausprägungen *rund*, *oval*, *lobuliert* und *irregulär* sind die im BI-RADS-Lexikon aufgeführten Merkmalsausprägungen für die Formbeschreibung von Raumforderungen, wobei *rund* und *oval* in dieser Arbeit einer gemeinsamen Kategorie zugeordnet wurden. Runde, ovale und lobulierte Herdbefunde sollen eher typisch für benigne Prozesse sein, während eine irreguläre Form zur Malignität tendiert (Nunes et al. 1997b, Baum et al. 2002, Liberman et al. 2002). In der Literatur finden sich auch weitere Merkmalszuweisungen im Sinne einer linearen, sternförmigen oder retikulären Form (Baum et al. 2002). Von sternförmigen Herdbeschreibungen ist bei den aktuell gängigen Publikationen kaum noch die Rede. Zurückzuführen ist dies eventuell auf die früher gebräuchliche unscharfe Trennung zwischen der Form- und Randbeurteilung. Die teils übliche Beschreibung einer sternförmigen Anreicherung könnte aktuell einer spikulierten Randbezeichnung gewichen sein. So unternahmen Buadu et al. eine Unterteilung der Randcharakterisierung in *glatt*, *lobuliert*, *irregulär* und *sternförmig* („*spiculated*“) (Buadu et al. 1996). Lineare Anreicherungsmuster werden heute meist der non-mass-Beschreibung zugewiesen. Ein retikuläres bzw. dendritisches Enhancement fällt ebenfalls in die non-mass-Domäne (Thomassin-Naggara et al. 2009).

Insgesamt war die Formbeschreibung in vorliegender Arbeit ein unsicheres Kriterium. Eine irreguläre Form war bei 59% der Malignome auszumachen, bei den gutartigen Befunden zu 31%. Fast umgekehrt war die Verteilung bei runden bzw. ovalen Läsionen. Hier waren 36% der bösartigen Läsionen rund/oval, bei den Gutartigen 54%. Lobulierte Befunde waren eher selten, unter den gutartigen Befunden stellten sich 15% dementsprechend dar, bei den malignen Anreicherungen 5%.

6.4 Vergleich mit anderen publizierten MRM-Interpretationsmodellen

Die mit 83,8% höchste diagnostische Genauigkeit dieser Arbeit wurde unter Einschluss sämtlicher evaluierter Deskriptoren erreicht, entsprechend einer AUC von 0,91. Durch Vorwärts- oder Rückwärtsselektion gelang die Reduktion auf zehn bzw. zwölf Variablen, die Modellgüte verschlechterte sich hierbei nicht signifikant.

Zwecks Übersichtlichkeit wurde die Anzahl der ins Modell aufzunehmenden Parameter auf fünf begrenzt. Beim Modell mit den Deskriptoren *zeitliches* bzw. *räumliches KM-Anreicherungsverhalten*, *Root-Zeichen*, *Ödem* und *T2-Helligkeit* verringerte sich die Genauigkeit signifikant auf 81,5% (AUC 0,89), bei Verwendung derselben kinetischen Deskriptoren, *Rand*, *Form* und *Septen* auf 80,4% (AUC 0,88). Letzteres Modell unterschied sich von allen anderen Modellen signifikant.

Zahlreiche Interpretationsmodelle wurden in der Literatur beschrieben. Die häufig geringe Fallzahl mit konsekutiv eingeschränkter allgemeiner Aussagekraft der Ergebnisse erfordert im Rahmen der vergleichenden Darstellung eine Beschränkung auf repräsentative Arbeiten.

Nunes et al. veröffentlichten ein erstes MRM-Interpretationsmodell, das lediglich auf morphologischen Beurteilungskriterien basierte und eine diagnostische Genauigkeit von 86% verzeichnen konnte (Nunes et al. 1997a). Eingeschlossen wurden hier auch Fälle ohne MR-mammographische Darstellung, denen jedoch ein Tastbefund oder eine mammographische Auffälligkeit vorausging. An dieser Studie nahmen 192 Patientinnen teil, die sich nach einer suspekten Mammographie oder einem Tastbefund einer MRM unterzogen. In 38 Fällen kam es zu keinem MR-mammographischen Befund. Rechnet man zum Zweck der Vergleichbarkeit mit der vorliegenden Arbeit die Fälle ohne Anreicherungen heraus, beträgt die diagnostische Genauigkeit 85,7%. Andererseits wurden Enhancements ohne palpatorische oder mammographische Entsprechung nicht berücksichtigt. Ebenso wurden Läsionen ausgeschlossen, welche zusätzlich zu anderen Anreicherungen auftraten, die nicht einem Tast- oder Mammographiebefund entsprachen. Überdies wurden hier auch MR-mammographisch eindeutig zuordenbare Zysten berücksichtigt, was nicht der Handhabung der vorliegenden und anderer Arbeiten entspricht, welche sich mit der Differenzierung anreichernder Herdbefunde befassen.

Der Ausschluss von Fällen, die in anderen Bildgebungsverfahren okkult sind, scheint die Genauigkeit eines Modells zu verbessern, da es sich oft um kleine Befunde handeln könnte,

die schwieriger zu beurteilen sind. Siegmann et al. haben eine Studie mit ausschließlich in der MRM sichtbaren Befunden durchgeführt und die im Vergleich zu anderen Arbeiten mit sonographischem oder mammographischem Korrelat schlechteren Ergebnisse mit dieser Argumentation begründet (Siegmann et al. 2009).

Weiterhin war mit 154 Läsionen die Fallzahl wesentlich geringer als diejenige, die der vorliegenden Arbeit zugrunde liegt. Es existierten außerdem einige sicher klassifizierbare Merkmale oder Merkmalskombinationen. Diese basierten allerdings auf geringen Fallzahlen, so dass eine allgemeingültige Aussage zurückhaltend getroffen werden sollte. Teilweise ergaben sich auch widersprüchliche Ergebnisse im Vergleich zu den Resultaten der vorliegenden Arbeit. Exemplarisch kann man hier die Annahme aufführen, dass eine glatt berandete Läsion immer gutartig gewesen sein sollte, was in der Nunes-Arbeit bei 21 von 154 Läsionen der Fall war. Wir fanden dagegen 273 scharf begrenzte Anreicherungen, von denen mit 194 nur 71% gutartig waren.

Ein von U. Fischer et al. vorgeschlagenes Scoring-Modell (Göttingen-Score) wurde von Baum und Mitarbeitern an einem größeren Kollektiv von 522 Patientinnen getestet, bei denen 265 Läsionen in 244 Mammæ beschrieben werden konnten (Fischer et al. 1999, Baum et al. 2002). Im Gegensatz zum Nunes-Modell wurden hier morphologische und kinetische Merkmale kombiniert. Die Dignitätsvorhersage wurde anhand der Ausprägung von fünf Deskriptoren vorgenommen, indem der initiale Signalanstieg, die Signalintensität nach der anfänglichen Signalanreicherung, die Form, die Berandung und das interne Anreicherungsmuster beurteilt wurden. Entsprechend des sich abhängig von den jeweiligen Merkmalsausprägungen ergebenden Scores ließ sich eine Diagnose ableiten. Baum et al. ermittelten bei dieser Verfahrensweise jeweils Werte von 92% für Sensitivität, Spezifität und diagnostische Genauigkeit, einer Anzahl an 243 korrekt klassifizierten von insgesamt 265 Anreicherungen entsprechend.

D. Fischer et al. fanden hingegen davon abweichende Ergebnisse (Fischer et al. 2005). 132 mit demselben Score evaluierte histologisch verifizierte Fälle ergaben deutlich niedrigere statistische Kennwerte mit einer Sensitivität von 81% und einer Spezifität von lediglich 59%. Zur Verbesserung des Modells wurde für die Bewertung von Befunden mit grenzwertigen Scorewerten eine Berücksichtigung zahlreicher Deskriptoren vorgeschlagen, die von Malich et al. als erweiterte Hilfsmittel für die Differenzierung maligner und benignen Befunde vorgeschlagen worden waren (Malich et al. 2005). Hierbei gelang es, die Sensitivität bzw.

Spezifität auf 97% bzw. 76,5% zu steigern, was sich auch in einer deutlich verbesserten diagnostischen Genauigkeit von 86,4% (114/132) zeigte. Viele der dort zusätzlich benutzten Deskriptoren wurden auch in vorliegender Arbeit verwendet, darunter das Root-, Hook- und Blooming-Zeichen, die Signalintensität in den T2-Aufnahmen, das Vorhandensein von Ödemen, Cutisverdickung, Mamillensaumunterbrechung, zur Läsion führende Gefäße und die Binnenarchitektur in nativen T1-gewichteten Bildern.

Form, Berandung, Binnenanreicherungsmuster und postinitiale Kinetik führten bei einem von Tozaki et al. vorgeschlagenen Modell zu einer Sensitivität von 99%, einer Spezifität von 89% und einem PPV bzw. NPV von 96% bzw. 98%, was eine diagnostische Genauigkeit von deutlich über 90% ergibt (Tozaki et al. 2005). Hier ist ein Zustandekommen der guten Ergebnisse ebenfalls kritisch zu hinterfragen. Zum einen wiesen auch hier bestimmte Zeichen wie bei Nunes 100%ige prädiktive Werte auf, was durch gängige Publikationen nicht bestätigt werden konnte. Dies betraf u. a. die spikulierten Herdbefunde, welche immer maligne sein sollten. In verschiedenen Veröffentlichungen wurde jedoch auch von geringeren Werten berichtet, die das Vorliegen eines Malignoms in beispielsweise nur 76% der Fälle mit spikuliertem Rand angeben (Nunes et al. 1997b). Andererseits ergaben sich auch hier aus der relativ kleinen Fallanzahl von 171 Läsionen bestimmte Befunde mit Merkmalskombinationen, anhand deren Anzahl nicht von einer allgemeinen Gültigkeit ausgegangen werden sollte. Ein homogen anreichernder Herd mit regelmäßiger Morphologie (rund, oval oder lobuliert) und Wash-out-Kinetik wäre beispielsweise immer benigne, allerdings wurde dies in jener Veröffentlichung lediglich viermal beobachtet.

Ein von Schnall und Kollegen publiziertes Interpretationsmodell nutzte ebenfalls die binäre logistische Regression zur Malignitätsabschätzung (Schnall et al. 2006). Man bediente sich hierbei der üblichen Deskriptoren wie Berandung, räumliches und zeitliches Verhalten der KM-Anreicherung. Allerdings wurde der MR-Befund nicht lediglich anhand morphologischer und kinetischer Läsionskriterien beurteilt. Auch das Lebensalter der Patientin, welches ja Hauptrisikofaktor für ein Mammamalignom ist, suspekte Mammographiebefunde mit Mikrokalzifikationen in der Vorgeschichte und das Vorhandensein eines palpablen Befundes gingen in die Modellberechnung ein. Es resultierte mit 0,88 ein AUC-Wert, der aufzeigt, dass trotz der zusätzlich verwendeten Modellparameter kein Vorteil gegenüber den Modellen dieser Arbeit nachgewiesen werden konnte.

6.5 Publikationsvergleich hinsichtlich einer sicheren Dignitätszuweisung

Die Eignung der Klassifikationsmodelle bezüglich der Filterung von Befunden, bei denen man sich hinsichtlich der diagnostischen Festlegung sicher ist, kann anhand von drei Aspekten beurteilt werden:

- 1) Wie hoch ist der Anteil der mutmaßlich sicher klassifizierten Befunde an allen betrachteten Anreicherungen?
- 2) Wie hoch ist hierbei die Rate an Fällen, bei denen man trotzdem falsch liegt?
- 3) Ist das jeweilige Modell besser zur Detektion maligner oder benigner Läsionen geeignet bzw. gelingt beides vergleichbar gut?

Zieht man die Daten von Baum et al. zurate, so lässt sich hier die beste Zuordnung zu sicher malignen oder benignen Befunden ausmachen. Dies kann anhand Zuweisung der Scores zu einer BI-RADS-Kategorie nachvollzogen werden. Desto höher der Punktwert ist, umso mehr steigt das Malignitätsrisiko. Die BI-RADS-Kategorien bilden eine Ordinalskala. Bei BI-RADS 1 findet sich ein unauffälliger MR-Befund, bei BI-RADS 2 geht man mit größter Wahrscheinlichkeit von einem benignen und bei BI-RADS 5 mit vergleichbarer Sicherheit von einem malignen Prozess aus. In der Konsequenz ergab sich eine in 59% aller Fälle sichere Zuordnung, die Quote hierbei falsch bewerteter Untersuchungen betrug 3,2%. 85 von 134 Malignomen wurden als sicher bösartig kategorisiert (63,4%), 72 von 131 sollten mit hoher Wahrscheinlichkeit gutartig sein (55%). In einem Fall lag man bei den vermeintlich sicheren Malignomen falsch (1,2%). Mit vier falschen Diagnosen bei den als sicher gutartig klassifizierten Veränderungen war die Fehlerrate höher (5,6%).

Auch hinsichtlich der sicheren Festlegung bezüglich der Dignität stehen den guten Ergebnissen von Baum et al. die Daten von D. Fischer et al. gegenüber, wenn man die Verfahrensweise zur Detektion sicher maligner bzw. benigner Befunde überträgt. Der Anteil falscher Befunde betrug sowohl bei den als sicher geglaubten Malignomen wie auch bei den gutartigen Veränderungen 9%, somit ist der Score mit wesentlich mehr Unsicherheiten behaftet, wenn man sich aufgrund bestimmter Merkmalskonstellationen mit größter Sicherheit auf die Dignität festlegen will.

Wie bereits erwähnt, verbesserte sich die diagnostische Performance bei Erweiterung durch die von Malich et al. eingeführten Deskriptoren. Trotzdem erfolgte eine Zuordnung zu den eine hohe diagnostische Sicherheit adressierenden BI-RADS-Kategorien 2 und 5 mit nicht akzeptablen Fehlerraten von 23% bzw. 18%.

Tozaki et al. veröffentlichten ebenfalls Merkmale oder Merkmalskombinationen, die mit 100%igen prädiktiven Werten aufwarteten. Dies war bei spikulierten Herdberandungen und einem nach dem Verständnis dieser Arbeit zentripetalem Enhancement (delayed central enhancement) der Fall. Anreichernde Septen (PPV=97%) und irreguläre Form (PPV=97%) erreichten fast selbige Werte. Mit den Ergebnissen dieser Arbeit und solchen anderer Publikationen sind diese Befunde jedoch nicht in Einklang zu bringen, was bei der Diskussion der Deskriptoren bereits erläutert wurde.

Die Nunes-Arbeit kann ausschließlich der Detektion von sicher benignen Befunden gerecht werden. So sind glatt berandete Läsionen zu 100% als gutartig zu betrachten, was sich jedoch im Vergleich zu dieser Arbeit widerspricht. Dies traf bei 21 von 154 Anreicherungen, hiervon 79 benignen, zu. Ein NPV von ebenfalls 100% wurde auch bei der Merkmalskombination einer lobuliert anreichernden Läsion mit Septen errechnet, was bei weiteren 20 Raumforderungen zutraf. Ein Befund war auch sicher benigne, wenn er zwar keine Septen aufwies, aber lobuliert war und nur schwach anreicherte (11/79) bzw. wenn die Herdbegrenzung spikuliert war und der Herd selbst nur schwach KM aufnahm (2/79) oder wenn, wie es in einem Fall geschah, die Begrenzung irregulär war und der Herd keine Septen aufwies (1/79). Somit wurde in 55 von 79 benignen Enhancements (70%) eine sichere Aussage bezüglich der Gutartigkeit getätigt. Eine relativ sichere Entscheidung zugunsten einer malignen Diagnose konnte prinzipiell nur getroffen werden, wenn der Herd ohne Septen und irregulär begrenzt war oder eine mäßige bis starke Anreicherung spikulierte Begrenzungen aufwies. Bei ersterer Kombination ergab sich eine Falsch-Positiv-Rate von 15% (4/26), bei letzterer 7% (2/30). Die in dieser Arbeit geforderte Sicherheit von 95% für Malignität, einhergehend mit einer sich daraus ergebenden Falsch-Positiv-Rate von max. 5% konnte mit diesem Modell also für Malignome nicht erfüllt werden. Somit ist dieses Interpretationsmodell nicht geeignet, Malignome mit an Sicherheit grenzender Wahrscheinlichkeit zu detektieren, die guten Ergebnisse hinsichtlich der sicheren Erkennung gutartiger Mammaveränderungen sind kritisch zu betrachten, einerseits weil die Fallzahlen für bestimmte Merkmalskombinationen zu gering

sind, andererseits weil die hier gefundenen Ergebnisse mit vorliegender Arbeit und mit anderen Publikationen im Widerspruch stehen.

Die durch Schnall et al. publizierten Daten ließen keine Schlussfolgerungen bezüglich einer sicheren Detektion maligner oder benigner Läsionen zu (Schnall et al. 2006).

Aufgrund des Vergleiches mit den hier aufgeführten Interpretationsmodellen lässt sich feststellen, dass die durch diese Arbeit errechneten binär-logistischen Regressionsmodelle geeignet sind, einen gewissen Anteil an Läsionen zu selektieren, bei denen mit höchster Sicherheit ein maligner oder benigner histologischer Befund zu erwarten wäre. In Abhängigkeit der verwendeten Deskriptoren liegt man in lediglich ca. 1-2% der Fälle falsch. Somit wird für bestimmte Fälle eine diagnostische Sicherheit erreicht, die derjenigen der histologischen Aufarbeitung nicht nachsteht. Beschränkt man sich auf eine geringer gehaltene übersichtlichere Anzahl an Deskriptoren, so sollte man zur Bestimmung sicher maligner und sicher benigner Befunde unterschiedliche binäre logistische Modelle anwenden.

6.6 Ausblick

Aufgrund zahlreicher in der Diskussion aufgeführter Aspekte lassen sich mehrere Ansätze für zukünftige Arbeiten zur Klassifikation von Herdbefunden der Mamma ableiten. Die Genauigkeit der hier vorgeschlagenen Modelle war limitiert. Auch wenn sich ein erheblicher Teil der untersuchten Läsionen mit einer hohen Genauigkeit klassifizieren ließ, ist die Treffsicherheit für eine Routineanwendung verbesserungswürdig. Dies liegt zum einen an der Fallselektion. Die Wahl der Histopathologie als Referenzstandard schließt bei bekannt hohem negativem Vorhersagewert der MRM einen Großteil der richtig-negativen Befunde aus und führt so methodisch bedingt zu einer Unterschätzung der Spezifität der Methode. Ebenso ist bei fehlender Verlaufsbetrachtung potentiell die Möglichkeit übersehener falsch-negativer Karzinome gegeben. Letztere ist aufgrund der hohen Sensitivität der MRM nicht in höherem Ausmaß zu erwarten. Die hier verwendeten Deskriptoren berücksichtigten - wie diskutiert - nicht die neuere amerikanische Einteilung in solide und nicht solide Befunde (ACR 2003). Eine solche Differenzierung wird potentiell der biologischen Situation gerechter und sollte demnach systematisch adressiert werden. Die vorliegende Studie war eine Langzeituntersuchung mit hoher Fallzahl und begann vor der Publikation des BI-RADS-Protokolls.

Die in dieser Arbeit vorliegende hohe Fallzahl ist für die reliable Erstellung von Regressionsmodellen nach Bland erforderlich (Bland 2000). Es wurde eine Vielzahl von Deskriptoren analysiert, welche gemäß der Regressionsanalyse z.T. redundante Informationen beinhalten. Auch wenn kritisch angemerkt werden muß, dass klassische statistische Verfahren wie die binäre logistische Regression den Einzelfall rechnerisch bedingt vernachlässigen, sollte dennoch eine Präzisierung morphologischer und kinetischer Deskriptoren mit Abbau von Redundanzen erfolgen. Darüber hinaus ist die Adressierung der IOR der genannten Deskriptoren erforderlich. Nur reliabel zu beurteilende sowie hinsichtlich der Herddifferenzierung valide Deskriptoren kommen für klinisch anwendbare Modelle in Betracht.

In den letzten Jahren finden zunehmend neuere technische Ansätze wie Analyse des Kontrastmittelverhaltens mittels spezieller Software oder pharmakokinetische Modelle, die Messung diffusionsgewichteter Sequenzen sowie die MR-Spektroskopie Anwendung. Alle diese Verfahren lassen sich – mit jeweils spezifischen, mehr oder weniger ausgeprägten Limitationen – quantifizieren (Williams et al. 2007, Mountford et al. 2009, Tsushima et al. 2009, Zwick et al. 2009). Eine quantitative Bildgebung birgt den Vorteil einer objektivierbaren Einschätzung von Pathologien. Eine Kombination solcher Techniken mit herkömmlicher morphologischer Beurteilung stellt somit einen innovativen Ansatz dar, welcher für die Kombination von diffusionsgewichteter und kontrastmittelgestützter Bildgebung bereits initial publiziert wurde (Yabuuchi et al. 2008, Yabuuchi et al. 2009).

7 Schlussfolgerungen

Die Bedeutung international etablierter morphologischer und dynamischer MR-Zeichen wie *Rand*, *Septen*, *postinitialer Signalintensitätsverlauf* und *KM-Anreicherungsverhalten* konnte durch die Ergebnisse der vorliegenden Arbeit im Wesentlichen bestätigt werden. Darüber hinaus gelang die Identifizierung zusätzlicher, bisher hauptsächlich am Universitätsklinikum Jena verwendeter Kriterien, welche für die Differentialdiagnose richtungsweisend sein können. Eine hohe Bedeutung nimmt dabei die Beurteilung auf Vorliegen des *Root*- und *Hook*-Zeichens ein, womit der Betrachtung von Herdausläufern eine wichtige Rolle zukommt. Ein *perifokales Ödem* macht eine maligne Diagnose ebenso sehr wahrscheinlich. Nicht zuletzt kann auch die Signalintensität in der T2-Aufnahme von Bedeutung sein. Während die isointense und hypointense Darstellung des Herdes wenig spezifisch ist, zeigt sich bei Hyperintensität eine starke Tendenz hin zu einem gutartigen Befund. Es existieren des Weiteren relativ seltene Befunde wie *Cutisverdickung* oder *Mamillensaumunterbrechung*. Treten sie jedoch auf, so spricht dies meist für einen malignen Prozess. Ebenfalls auf Malignität weist die vergleichsweise häufigere Prävalenz eines auf den Herdbefund zuführenden Gefäßes hin. Weniger bedeutend in ihrer Aussagekraft erscheint die Beurteilung nativer T1w-Bilder hinsichtlich Signalintensität und Binnenstruktur.

Im Rahmen dieser Arbeit wurden Regressionsmodelle mit hoher Genauigkeit entwickelt, welche zudem eine sichere Klassifikation bei über 95%iger Sicherheit von gut einem Viertel der Herdbefunde bei akzeptablen Fehlerraten (<2%) ermöglichten. Darüber hinaus konnten vereinfachte Modelle ohne wesentlichen Verlust an diagnostischer Genauigkeit erstellt werden. Dies ist als ein Hinweis auf die Machbarkeit einer klinischen Differentialdiagnose anhand weniger standardisierter Kriterien zu deuten. Die relativ unkomplizierte Deskriptorenselektion impliziert allerdings auch die Möglichkeit redundanter Information zwischen der Vielzahl aktuell existierender MR-Zeichen, zumindest vom Blickpunkt klassischer Statistik aus. Künftige Arbeiten könnten sich daher mit exakteren Deskriptoren befassen, darunter auch solchen, welche die unterschiedlichen anatomischen Wachstumsmuster von soliden Tumoren und non-mass-Läsionen berücksichtigen. Die in dieser Arbeit verwendeten Deskriptoren sind rein visuell-qualitativer Natur, ermöglichen jedoch die Erstellung valider Vorhersagemodelle. Die Reliabilität derselben wurde bisher nicht systematisch adressiert.

Quantitative Deskriptoren könnten die diagnostische Aussage der MRM potentiell objektivieren oder gar verbessern, ein systematischer Beweis für diese These müsste jedoch zuvor erbracht werden.

Zuletzt erfordert die umfassende Einführung von Interpretationsmodellen in die klinische Praxis prospektive Studien im klinischen Kontext, da bei vorliegender Arbeit lediglich histologisch verifizierte Befunde verwendet wurden und somit selektionsbedingt verzerrte statistische Parameter in Form einer niedrigeren Spezifität sowie eines höheren positiven Vorhersagewertes im Vergleich zur klinischen Wirklichkeit zu erwarten sind. Die in dieser Arbeit ermittelten diagnostischen Genauigkeiten beschreiben nicht die Treffsicherheit der Methode als solche, sondern lediglich mathematische Modelle subjektiver Deskriptoren in dem hier definierten Patientengut.

8 Literaturverzeichnis

- Albert US, Schulz KD. 2003. Selbstuntersuchung. In: K.-D. Schulz, U.-S. Albert und die Mitglieder der Planungsgruppe und Leiter der Arbeitsgruppen Konzertierte Aktion Brustkrebs-Früherkennung in Deutschland (Hrsg.). Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland. München · Wien · New York: W. Zuckschwerdt Verlag, S. 49-53.
- American College of Radiology. 2003. Breast imaging reporting and data system atlas [BI-RADS atlas]. Reston, VA: American College of Radiology.
- Altman DG, Bland JM. 1994. Diagnostic tests 2: Predictive values. *BMJ*, 309 (6947):102.
- Baltzer PA, Renz DM, Kullnig PE, Gajda M, Camara O, Kaiser WA. 2009a. Application of computer-aided diagnosis (CAD) in MR-mammography (MRM): do we really need whole lesion time curve distribution analysis? *Acad Radiol*, 16 (4):435-442.
- Baltzer PA, Benndorf M, Dietzel M, Gajda M, Camara O, Kaiser WA. 2009b. Sensitivity and specificity of unenhanced MR mammography (DWI combined with T2-weighted TSE imaging, ueMRM) for the differentiation of mass lesions. *Eur Radiol*.
- Baltzer PA, Freiberg C, Beger S, Vag T, Dietzel M, Herzog AB, Gajda M, Camara O, Kaiser WA. 2009c. Clinical MR-mammography: are computer-assisted methods superior to visual or manual measurements for curve type analysis? A systematic approach. *Acad Radiol*, 16 (9):1070-1076.
- Bässler R. 1997 Mamma. In: Remmele W (Hrsg.): Pathologie 4: Weibliches Genitale – Mamma – Pathologie der Schwangerschaft, der Plazenta und des Neugeborenen – Infektionskrankheiten des Fetus und des Neugeborenen, Tumoren des Kindesalters – Endokrine Organe. Zweite, neu bearbeitete Auflage. Berlin · Heidelberg · New York: Springer, S. 135-365.

- Batzler WU. 2006. Krebs in Deutschland. Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. und das RKI (Hrsg.) 5. überarbeitete, aktualisierte Ausgabe. Saarbrücken.
- Batzler WU, Giersiepen K, Hentschel S, Husmann G, Kaatsch P, Katalinic A, Kieschke J, Kraywinkel K, Meyer M, Stabenow R, Stegmaier C, Bertz J, Haberland J, Wolf U. 2008. Brustdrüse der Frau. In: Robert Koch-Institut (Hrsg.) und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e. V. (Hrsg.). Krebs in Deutschland 2003 – 2004. Häufigkeiten und Trends. 6. überarbeitete Auflage. Berlin, S. 54-57.
- Baum F, Fischer U, Vosschenrich R, Grabbe E. 2002. Classification of hypervascularized lesions in CE MR imaging of the breast. Eur Radiol, 12 (5):1087-1092.
- Beral V. 2003. Breast cancer and hormone-replacement therapy in the Million Women Study. Lancet, 362 (9382):419-427.
- Bick U. 2000. [Typical and unusual findings in MR mammography]. Rofo, 172 (5):415-428.
- Bland M. 2000. An Introduction to Medical Statistics. Oxford, UK: Oxford University Press.
- Bloch F, Hansen WW and Packard M. 1946. The Nuclear Induction Experiment. Phys Rev 70: 474.
- Bock K, Albert US, Schulz KD. 2003 Hormone und Brustdrüsengewebsdichte. In: K.-D. Schulz, U.-S. Albert und die Mitglieder der Planungsgruppe und Leiter der Arbeitsgruppen Konzertierte Aktion Brustkrebs-Früherkennung in Deutschland (Hrsg.). Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland. München · Wien · New York: W. Zuckschwerdt Verlag, S. 68-73.
- Boetes C, Barentsz JO, Mus RD, van der Sluis RF, van Erning LJ, Hendriks JH, Holland R, Ruys SH. 1994. MR characterization of suspicious breast lesions with a gadolinium-enhanced TurboFLASH subtraction technique. Radiology, 193 (3):777-781.

- Bone B, Aspelin P, Isberg B, Perbeck L, Veress B. 1995. Contrast-enhanced MR imaging of the breast in patients with breast implants after cancer surgery. *Acta Radiol*, 36 (2):111-116.
- Bray F, Sankila R, Ferlay J, Parkin DM. 2002. Estimates of cancer incidence and mortality in Europe in 1995. *Eur J Cancer*, 38 (1):99-166.
- Breiman L, Friedman J, Olshen R, Stone C. 1984. Classification and regression trees. Wadsworth, CA: Pacific Grove.
- Buadu LD, Murakami J, Murayama S, Hashiguchi N, Sakai S, Masuda K, Toyoshima S, Kuroki S, Ohno S. 1996. Breast lesions: correlation of contrast medium enhancement patterns on MR images with histopathologic findings and tumor angiogenesis. *Radiology*, 200 (3):639-649.
- Buadu LD, Murakami J, Murayama S, Hashiguchi N, Sakai S, Toyoshima S, Masuda K, Kuroki S, Ohno S. 1997. Patterns of peripheral enhancement in breast masses: correlation of findings on contrast medium enhanced MRI with histologic features and tumor angiogenesis. *J Comput Assist Tomogr*, 21 (3):421-430.
- Carmichael AR, Harbach L, Cooke R. 2009. Breast clinic and life style study BLLISS. *Int Semin Surg Oncol*, 6:12.
- Carr DH, Brown J, Bydder GM, Steiner RE, Weinmann HJ, Speck U, Hall AS, Young IR. 1984. Gadolinium-DTPA as a contrast agent in MRI: initial clinical experience in 20 patients. *AJR Am J Roentgenol*, 143 (2):215-224.
- Collaborative Group. 1996. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. Collaborative Group on Hormonal Factors in Breast Cancer. *Lancet*, 347 (9017):1713-1727.

- Crystal P, Koretz M, Shcharynsky S, Makarov V, Strano S. 2005. Accuracy of sonographically guided 14-gauge core-needle biopsy: results of 715 consecutive breast biopsies with at least two-year follow-up of benign lesions. *J Clin Ultrasound*, 33 (2):47-52.
- Damadian R, Zaner K, Hor D, DiMaio T, Minkoff L, Goldsmith M. 1973. Nuclear magnetic resonance as a new tool in cancer research: human tumors by NMR. *Ann N Y Acad Sci*, 222:1048-1076.
- Deligeoroglou E, Michailidis E, Creatsas G. 2003. Oral contraceptives and reproductive system cancer. *Ann N Y Acad Sci*, 997:199-208.
- Dietzel M, Baltzer PA, Vag T, Gröschel T, Gajda M, Camara O, Kaiser WA. 2009. Differential diagnosis of breast lesions ≤ 5 mm. Is there a role for MRI? *J Comput Assist Tomogr* [accepted].
- Dvorak HF, Dickersin GR, Dvorak AM, Manseau EJ, Pyne K. 1981. Human breast carcinoma: fibrin deposits and desmoplasia. Inflammatory cell type and distribution. Microvasculature and infarction. *J Natl Cancer Inst*, 67 (2):335-345.
- El Yousef SJ, Duchesneau RH, Alfidi RJ, Haaga JR, Bryan PJ, LiPuma JP. 1984. Magnetic resonance imaging of the breast. Work in progress. *Radiology*, 150 (3):761-766.
- Erguvan-Dogan B, Whitman GJ, Kushwaha AC, Phelps MJ, Dempsey PJ. 2006. BI-RADS-MRI: a primer. *AJR Am J Roentgenol*, 187 (2):W152-160.
- Erren TC, Morfeld P, Stork J, Knauth P, von Mulmann MJ, Breitstadt R, Muller U, Emmerich M, Piekarski C. 2009. Shift work, chronodisruption and cancer?--The IARC 2007 challenge for research and prevention and 10 theses from the Cologne Colloquium 2008. *Scand J Work Environ Health*, 35 (1):74-79.

- Fahrbach K, Sledge I, Cella C, Linz H, Ross SD. 2006. A comparison of the accuracy of two minimally invasive breast biopsy methods: a systematic literature review and meta-analysis. *Arch Gynecol Obstet*, 274 (2):63-73.
- Fischer DR, Wurdinger S, Boettcher J, Malich A, Kaiser WA. 2005. Further signs in the evaluation of magnetic resonance mammography: a retrospective study. *Invest Radiol*, 40 (7):430-435.
- Fischer DR, Malich A, Wurdinger S, Boettcher J, Dietzel M, Kaiser WA. 2006. The adjacent vessel on dynamic contrast-enhanced breast MRI. *AJR Am J Roentgenol*, 187 (2):W147-151.
- Fischer DR, Baltzer P, Malich A, Wurdinger S, Freesmeyer MG, Marx C, Kaiser WA. 2004. Is the "blooming sign" a promising additional tool to determine malignancy in MR mammography? *Eur Radiol*, 14 (3):394-401.
- Fischer U, Kopka L, Grabbe E. 1999. Breast carcinoma: effect of preoperative contrast-enhanced MR imaging on the therapeutic approach. *Radiology*, 213 (3):881-888.
- Fischer U, Vosschenrich R, Kopka L, Kahlen O, Grabbe E. 1996. [Contrast medium assisted dynamic MR-mammography after diagnostic and therapeutic interventions on the breast]. *Bildgebung*, 63 (2):94-100.
- Fobben ES, Rubin CZ, Kalisher L, Dembner AG, Seltzer MH, Santoro EJ. 1995. Breast MR imaging with commercially available techniques: radiologic-pathologic correlation. *Radiology*, 196 (1):143-152.
- Folkman J. 1992. The role of angiogenesis in tumor growth. *Semin Cancer Biol*, 3 (2):65-71.
- Folkman J, Klagsbrun M. 1987. Angiogenic factors. *Science*, 235 (4787):442-447.
- Folkman J, Watson K, Ingber D, Hanahan D. 1989. Induction of angiogenesis during the transition from hyperplasia to neoplasia. *Nature*, 339 (6219):58-61.

- Gray S. 2003. Breast cancer and hormone-replacement therapy: the Million Women Study. *Lancet*, 362 (9392):1332; author reply 1332.
- Guth U, Huang DJ, Huber M, Schotzau A, Wruk D, Holzgreve W, Wight E, Zanetti-Dallenbach R. 2008. Tumor size and detection in breast cancer: Self-examination and clinical breast examination are at their limit. *Cancer Detect Prev*, 32 (3):224-228.
- Haase A, Matthaei D, Hanicke W, Frahm J. 1986. Dynamic digital subtraction imaging using fast low-angle shot MR movie sequence. *Radiology*, 160 (2):537-541.
- Harms SE, Flamig DP, Hesley KL, Meiches MD, Jensen RA, Evans WP, Savino DA, Wells RV. 1993a. MR imaging of the breast with rotating delivery of excitation off resonance: clinical experience with pathologic correlation. *Radiology*, 187 (2):493-501.
- Harms SE, Flamig DP, Hesley KL, Evans WP, 3rd, Cheek JH, Peters GN, Knox SM, Savino DA, Netto GJ, Wells RB, et al. 1993b. Fat-suppressed three-dimensional MR imaging of the breast. *Radiographics*, 13 (2):247-267.
- Harris JR, Lippman ME, Veronesi U, Willett W. 1992. Breast cancer (1). *N Engl J Med*, 327 (5):319-328.
- Heiberg EV, Perman WH, Herrmann VM, Janney CG. 1996. Dynamic sequential 3D gadolinium-enhanced MRI of the whole breast. *Magn Reson Imaging*, 14 (4):337-348.
- Hermann G, Janus C, Schwartz IS, Papatestas A, Hermann DG, Rabinowitz JG. 1988. Occult malignant breast lesions in 114 patients: relationship to age and the presence of microcalcifications. *Radiology*, 169 (2):321-324.
- Heywang-Kobrunner SH, Haustein J, Pohl C, Beck R, Lommatzsch B, Untch M, Nathrath WB. 1994. Contrast-enhanced MR imaging of the breast: comparison of two different doses of gadopentetate dimeglumine. *Radiology*, 191 (3):639-646.

- Heywang-Köbrunner SH, Altland H. 2003. Magnetresonanztomographie. In: K.-D. Schulz, U.-S. Albert und die Mitglieder der Planungsgruppe und Leiter der Arbeitsgruppen Konzierte Aktion Brustkrebs-Früherkennung in Deutschland (Hrsg.). Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland. München · Wien · New York: W. Zuckschwerdt Verlag, S. 74-77.
- Hochman MG, Orel SG, Powell CM, Schnall MD, Reynolds CA, White LN. 1997. Fibroadenomas: MR imaging appearances with radiologic-histopathologic correlation. *Radiology*, 204 (1):123-129.
- Holland BA, Brant-Zawadzki M, Norman D, Newton TH. 1985. Magnetic resonance imaging of primary intracranial tumors: a review. *Int J Radiat Oncol Biol Phys*, 11 (2):315-321.
- Ikeda DM, Hylton NM, Kinkel K, Hochman MG, Kuhl CK, Kaiser WA, Weinreb JC, Smazal SF, Degani H, Viehweg P, Barclay J, Schnall MD. 2001. Development, standardization, and testing of a lexicon for reporting contrast-enhanced breast magnetic resonance imaging studies. *J Magn Reson Imaging*, 13 (6):889-895.
- Jackson VP. 1995. The current role of ultrasonography in breast imaging. *Radiol Clin North Am*, 33 (6):1161-1170.
- Jansen SA, Fan X, Karczmar GS, Abe H, Schmidt RA, Giger M, Newstead GM. 2008. DCEMRI of breast lesions: is kinetic analysis equally effective for both mass and nonmass-like enhancement? *Med Phys*, 35 (7):3102-3109.
- Kaiser W, Zeitler E. 1984. [Digital-subtracted magnetic resonance (DSMR). A new procedure for improved imaging of vascular processes]. *Rofo*, 141 (5):524-528.
- Kaiser WA. 1989. [Magnetic resonance tomography of the breast. The results of 253 examinations]. *Dtsch Med Wochenschr*, 114 (36):1351-1357.
- Kaiser WA. 1992. MRM promises earlier breast cancer diagnosis. *Diagn Imaging (San Franc)*, 14 (9):88-93.

Kaiser WA. 1993. [MR mammography]. Radiologe, 33 (5):292-299.

Kaiser WA. 2007. Signs in MR-Mammography. Erste Auflage. Berlin: Springer.

Kaiser WA, Zeitler E. 1989. MR imaging of the breast: fast imaging sequences with and without Gd-DTPA. Preliminary observations. Radiology, 170 (3 Pt 1):681-686.

Kaiser WA, Kess H. 1989. [A prototype double coil for MR breast measurement]. Rofo, 151 (1):103-105.

Kalayanarooj S. 2008. Benign and malignant soft tissue mass: magnetic resonance imaging criteria for discrimination. J Med Assoc Thai, 91 (1):74-81.

Kawashima M, Tamaki Y, Nonaka T, Higuchi K, Kimura M, Koida T, Yanagita Y, Sugihara S. 2002. MR imaging of mucinous carcinoma of the breast. AJR Am J Roentgenol, 179 (1):179-183.

Kelsey JL, Gammon MD, John EM. 1993. Reproductive factors and breast cancer. Epidemiol Rev, 15 (1):36-47.

Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. 1996. Likelihood ratios for modern screening mammography. Risk of breast cancer based on age and mammographic interpretation. JAMA, 276 (1):39-43.

Kinkel K, Helbich TH, Esserman LJ, Barclay J, Schwerin EH, Sickles EA, Hylton NM. 2000. Dynamic high-spatial-resolution MR imaging of suspicious breast lesions: diagnostic criteria and interobserver variability. AJR Am J Roentgenol, 175 (1):35-43.

Knopp MV, Bourne MW, Sardanelli F, Wasser MN, Bonomo L, Boetes C, Muller-Schimpfle M, Hall-Craggs MA, Hamm B, Orlacchio A, Bartolozzi C, Kessler M, Fischer U, Schneider G, Oudkerk M, Teh WL, Gehl HB, Salerio I, Pirovano G, La Noce A, Kirchin MA, Spinazzi A. 2003. Gadobenate dimeglumine-enhanced MRI of the breast: analysis of dose response and comparison with gadopentetate dimeglumine. *AJR Am J Roentgenol*, 181 (3):663-676.

Krebsregister Saarland: http://www.krebsregister.saarland.de/pdf/brustkrebs_erg.pdf

Kreienberg R. 2003. Klinische Untersuchung. In: K.-D. Schulz, U.-S. Albert und die Mitglieder der Planungsgruppe und Leiter der Arbeitsgruppen Konzertierte Aktion Brustkrebs-Früherkennung in Deutschland (Hrsg.). Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland. München · Wien · New York: W. Zuckschwerdt Verlag, S. 53-55.

Kuhl C. 2007a. The current status of breast MR imaging. Part I. Choice of technique, image interpretation, diagnostic accuracy, and transfer to clinical practice. *Radiology*, 244 (2):356-378.

Kuhl CK. 2007b. Breast MR imaging at 3T. *Magn Reson Imaging Clin N Am*, 15 (3):315-320, vi.

Kuhl CK, Klaschik S, Mielcarek P, Gieseke J, Wardelmann E, Schild HH. 1999a. Do T2-weighted pulse sequences help with the differential diagnosis of enhancing lesions in dynamic breast MRI? *J Magn Reson Imaging*, 9 (2):187-196.

Kuhl CK, Jost P, Morakkabati N, Zivanovic O, Schild HH, Gieseke J. 2006. Contrast-enhanced MR imaging of the breast at 3.0 and 1.5 T in the same patients: initial experience. *Radiology*, 239 (3):666-676.

- Kuhl CK, Mielcareck P, Klaschik S, Leutner C, Wardelmann E, Gieseke J, Schild HH. 1999b. Dynamic breast MR imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions? *Radiology*, 211 (1):101-110.
- Landis JR, Koch GG. 1977. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, 33 (2):363-374.
- Lanitis S, Rice AJ, Vaughan A, Cathcart P, Filippakis G, Al Mufti R, Hadjiminis DJ. 2008. Diagnosis and management of male breast cancer. *World J Surg*, 32 (11):2471-2476.
- Laya MB, Larson EB, Taplin SH, White E. 1996. Effect of estrogen replacement therapy on the specificity and sensitivity of screening mammography. *J Natl Cancer Inst*, 88 (10):643-649.
- Liberman L, Abramson AF, Squires FB, Glassman JR, Morris EA, Dershaw DD. 1998. The breast imaging reporting and data system: positive predictive value of mammographic features and final assessment categories. *AJR Am J Roentgenol*, 171 (1):35-40.
- Liberman L, Morris EA, Lee MJ, Kaplan JB, LaTrenta LR, Menell JH, Abramson AF, Dashnaw SM, Ballon DJ, Dershaw DD. 2002. Breast lesions detected on MR imaging: features and positive predictive value. *AJR Am J Roentgenol*, 179 (1):171-178.
- Madjar H, Duda VF, Schreer I. 2003. Mammasonographie. In: K.-D. Schulz, U.-S. Albert und die Mitglieder der Planungsgruppe und Leiter der Arbeitsgruppen Konzertierte Aktion Brustkrebs-Früherkennung in Deutschland (Hrsg.). *Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland*. München · Wien · New York: W. Zuckschwerdt Verlag, S. 73-74.
- Malich A, Fischer DR, Wurdinger S, Boettcher J, Marx C, Facius M, Kaiser WA. 2005. Potential MRI interpretation model: differentiation of benign from malignant breast masses. *AJR Am J Roentgenol*, 185 (4):964-970.

- Mann RM, Kuhl CK, Kinkel K, Boetes C. 2008. Breast MRI: guidelines from the European Society of Breast Imaging. *Eur Radiol*, 18 (7):1307-1318.
- Mansfield P, Morris PG, Ordidge R, Coupland RE, Bishop HM, Blamey RW. 1979. Carcinoma of the breast imaged by nuclear magnetic resonance (NMR). *Br J Radiol*, 52 (615):242-243.
- Morakkabati N, Leutner CC, Schmiedel A, Schild HH, Kuhl CK. 2003. Breast MR imaging during or soon after radiation therapy. *Radiology*, 229 (3):893-901.
- Mountford C, Ramadan S, Stanwell P, Malycha P. 2009. Proton MRS of the breast in the clinical setting. *NMR Biomed*, 22 (1):54-64.
- Müller-Schimpfle M, Ohmenhauser K, Claussen CD. 1997. [Effect of age and menstrual cycle on mammography and MR mammography]. *Radiologe*, 37 (9):718-725.
- Mussurakis S, Buckley DL, Coady AM, Turnbull LW, Horsman A. 1996. Observer variability in the interpretation of contrast enhanced MRI of the breast. *Br J Radiol*, 69 (827):1009-1016.
- Nunes LW, Schnall MD, Orel SG. 2001. Update of breast MR imaging architectural interpretation model. *Radiology*, 219 (2):484-494.
- Nunes LW, Schnall MD, Orel SG, Hochman MG, Langlotz CP, Reynolds CA, Torosian MH. 1997a. Breast MR imaging: interpretation model. *Radiology*, 202 (3):833-841.
- Nunes LW, Schnall MD, Siegelman ES, Langlotz CP, Orel SG, Sullivan D, Muenz LA, Reynolds CA, Torosian MH. 1997b. Diagnostic performance characteristics of architectural features revealed by high spatial-resolution MR imaging of the breast. *AJR Am J Roentgenol*, 169 (2):409-415.

- Nystrom L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S, Andersson I, Bjurstam N, Fagerberg G, Frisell J, et al. 1993. Breast cancer screening with mammography: overview of Swedish randomised trials. *Lancet*, 341 (8851):973-978.
- Obenauer S, Schorn C, Stelter B, Fischer U, Grabbe E. 2002. Contrast-enhanced high in-plane resolution dynamic MRI of the breast. Are there advantages in comparison to standard dynamic MRI? *Clin Imaging*, 26 (3):161-165.
- Pabst T, Kenn W, Kaiser WA, Hahn D. 2001. Understanding why contrast enhancement in dynamic MRI is not reproducible: illustration with a simple phantom. *Breast J*, 7 (3):166-170.
- Parkin DM, Bray F, Ferlay J, Pisani P. 2005. Global cancer statistics, 2002. *CA Cancer J Clin*, 55 (2):74-108.
- Peters NH, Borel Rinkes IH, Zuithoff NP, Mali WP, Moons KG, Peeters PH. 2008. Meta-analysis of MR imaging in the diagnosis of breast lesions. *Radiology*, 246 (1):116-124.
- Pike MC. 1987. Age-related factors in cancers of the breast, ovary, and endometrium. *J Chronic Dis*, 40 Suppl 2:59S-69S.
- Purcell EM, Torrey HC and Pound RV. 1946. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys Rev* 69: 37-43.
- Qayyum A, Birdwell RL, Daniel BL, Nowels KW, Jeffrey SS, Agoston TA, Herfkens RJ. 2002. MR imaging features of infiltrating lobular carcinoma of the breast: histopathologic correlation. *AJR Am J Roentgenol*, 178 (5):1227-1232.
- Reiser M, Kuhn FP, Debus J. 2006. Radiologie. Zweite, korrigierte Auflage. Stuttgart: Thieme.

- Renz DM, Baltzer PA, Bottcher J, Thaher F, Gajda M, Camara O, Runnebaum IB, Kaiser WA. 2008. Magnetic resonance imaging of inflammatory breast carcinoma and acute mastitis. A comparative study. *Eur Radiol*, 18 (11):2370-2380.
- Rieber A, Zeitler H, Rosenthal H, Gorich J, Kreienberg R, Brambs HJ, Tomczak R. 1997. MRI of breast cancer: influence of chemotherapy on sensitivity. *Br J Radiol*, 70 (833):452-458.
- Rissanen T, Pamilo M, Suramo I. 1998. Ultrasonography as a guidance method in the evaluation of mammographically detected nonpalpable breast lesions of suspected malignancy. *Acta Radiol*, 39 (3):292-297.
- Rosenberg RD, Hunt WC, Williamson MR, Gilliland FD, Wiest PW, Kelsey CA, Key CR, Linver MN. 1998. Effects of age, breast density, ethnicity, and estrogen replacement therapy on screening mammographic sensitivity and cancer stage at diagnosis: review of 183,134 screening mammograms in Albuquerque, New Mexico. *Radiology*, 209 (2):511-518.
- Ross RJ, Thompson JS, Kim K, Bailey RA. 1982. Nuclear magnetic resonance imaging and evaluation of human breast tissue: preliminary clinical trials. *Radiology*, 143 (1):195-205.
- Saarenmaa I, Salminen T, Geiger U, Heikkinen P, Hyvarinen S, Isola J, Kataja V, Kokko ML, Kokko R, Kumpulainen E, Karkkainen A, Pakkanen J, Peltonen P, Piironen A, Salo A, Talviala ML, Haka M. 2001. The effect of age and density of the breast on the sensitivity of breast cancer diagnostic by mammography and ultasonography. *Breast Cancer Res Treat*, 67 (2):117-123.
- Sardanelli F, Iozzelli A, Fausto A, Carriero A, Kirchin MA. 2005. Gadobenate dimeglumine-enhanced MR imaging breast vascular maps: association between invasive cancer and ipsilateral increased vascularity. *Radiology*, 235 (3):791-797.

- Schleicher UM, Ammon J. 1998. Mode of breast cancer detection: a study from the German part of the Maas-Rhine-EUREGIO. *Eur J Cancer Prev*, 7 Suppl 1:S41-46.
- Schnall MD, Rosten S, Englander S, Orel SG, Nunes LW. 2001. A combined architectural and kinetic interpretation model for breast MR images. *Acad Radiol*, 8 (7):591-597.
- Schnall MD, Blume J, Bluemke DA, DeAngelis GA, DeBruhl N, Harms S, Heywang-Kobrunner SH, Hylton N, Kuhl CK, Pisano ED, Causer P, Schnitt SJ, Thickman D, Stelling CB, Weatherall PT, Lehman C, Gatsonis CA. 2006. Diagnostic architectural and dynamic features at breast MR imaging: multicenter study. *Radiology*, 238 (1):42-53.
- Schneider BP, Miller KD. 2005. Angiogenesis of breast cancer. *J Clin Oncol*, 23 (8):1782-1790.
- Schreer I, Engel J. 2003. Mammographie. In: K.-D. Schulz, U.-S. Albert und die Mitglieder der Planungsgruppe und Leiter der Arbeitsgruppen Konzertierte Aktion Brustkrebs-Früherkennung in Deutschland (Hrsg.). *Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland*. München · Wien · New York: W. Zuckschwerdt Verlag, S. 56-68.
- Schueller G, Jaromi S, Ponhold L, Fuchsjaeger M, Memarsadeghi M, Rudas M, Weber M, Liberman L, Helbich TH. 2008. US-guided 14-gauge core-needle breast biopsy: results of a validation study in 1352 cases. *Radiology*, 248 (2):406-413.
- Schulz-Wendtland R, Kramer S, Doinghaus K, Mitze M, Lang N. 1997. [Interventional techniques in breast diagnosis: ultrasound controlled puncture biopsy]. *Aktuelle Radiol*, 7 (1):30-34.
- Senger DR, Connolly DT, Van de Water L, Feder J, Dvorak HF. 1990. Purification and NH₂-terminal amino acid sequence of guinea pig tumor-secreted vascular permeability factor. *Cancer Res*, 50 (6):1774-1778.

- Siegmann KC, Moron HU, Baur A, Hahn M, Vogel U, Claussen CD, Bitzer M. 2009. [Diagnostic value of a breast MRI score for the prediction of malignancy of breast lesions detected solely with MRI]. *Rofo*, 181 (6):556-563.
- Stelling CB, Wang PC, Lieber A, Mattingly SS, Griffen WO, Powell DE. 1985. Prototype coil for magnetic resonance imaging of the female breast. Work in progress. *Radiology*, 154 (2):457-462.
- Stomper PC, Herman S, Klippenstein DL, Winston JS, Edge SB, Arredondo MA, Mazurchuk RV, Blumenson LE. 1995. Suspect breast lesions: findings at dynamic gadolinium-enhanced MR imaging correlated with mammographic and pathologic features. *Radiology*, 197 (2):387-395.
- Swets JA, Getty DJ, Pickett RM, D'Orsi CJ, Seltzer SE, McNeil BJ. 1991. Enhancing and evaluating diagnostic accuracy. *Med Decis Making*, 11 (1):9-18.
- Szabo BK, Aspelin P, Wiberg MK, Bone B. 2003. Dynamic MR imaging of the breast. Analysis of kinetic and morphologic diagnostic criteria. *Acta Radiol*, 44 (4):379-386.
- Tardivon AA, Guinebretiere JM, Dromain C, Vanel D. 2002. Imaging and management of nonpalpable lesions of the breast. *Eur J Radiol*, 42 (1):2-9.
- Teh W, Wilson AR. 1998. The role of ultrasound in breast cancer screening. A consensus statement by the European Group for Breast Cancer Screening. *Eur J Cancer*, 34 (4):449-450.
- Thomas DB, Gao DL, Ray RM, Wang WW, Allison CJ, Chen FL, Porter P, Hu YW, Zhao GL, Pan LD, Li W, Wu C, Coriaty Z, Evans I, Lin MG, Stalsberg H, Self SG. 2002. Randomized trial of breast self-examination in Shanghai: final results. *J Natl Cancer Inst*, 94 (19):1445-1457.

- Thomassin-Naggara I, Salem C, Darai E, Bazot M, Uzan S, Marsault C, Chopier J. 2009. [Non-masslike enhancement on breast MRI: interpretation pearls]. *J Radiol*, 90 (3 Pt 1):269-275.
- Thull DL, Vogel VG. 2004. Recognition and management of hereditary breast cancer syndromes. *Oncologist*, 9 (1):13-24.
- Tozaki M, Igarashi T, Fukuda K. 2006. Positive and negative predictive values of BI-RADS-MRI descriptors for focal breast masses. *Magn Reson Med Sci*, 5 (1):7-15.
- Tozaki M, Igarashi T, Matsushima S, Fukuda K. 2005. High-spatial-resolution MR imaging of focal breast masses: interpretation model based on kinetic and morphological parameters. *Radiat Med*, 23 (1):43-50.
- Tsushima Y, Takahashi-Taketomi A, Endo K. 2009. Magnetic resonance (MR) differential diagnosis of breast tumors using apparent diffusion coefficient (ADC) on 1.5-T. *J Magn Reson Imaging*, 30 (2):249-255.
- Walker RA. 2001. The complexities of breast cancer desmoplasia. *Breast Cancer Res*, 3 (3):143-145.
- Warren R, Ciatto S, Macaskill P, Black R, Houssami N. 2009. Technical aspects of breast MRI--do they affect outcomes? *Eur Radiol*, 19 (7):1629-1638.
- Wielopolski P, Schraa B, Oudkerk M. 2001. Gd-DTPA empfindliche Pulssequenzen. In: Hosten N, Lemke JA, Felix R, Heshiki A, Hricak H, und Chang KH (Hrsg.): Monographie Gadolinium-DTPA Magnevist. Berlin · Wien: Blackwell Wissenschaft, S. 13-26.
- Williams TC, DeMartini WB, Partridge SC, Peacock S, Lehman CD. 2007. Breast MR imaging: computer-aided evaluation program for discriminating benign from malignant lesions. *Radiology*, 244 (1):94-103.

- Wolfman NT, Williams RW, Wall BE, Moran PR, Karstaedt N. 1986. Design modification of dedicated MR breast coil. *J Comput Assist Tomogr*, 10 (5):893-895.
- Wu Y, Doi K, Metz CE, Asada N, Giger ML. 1993. Simulation studies of data classification by artificial neural networks: potential applications in medical imaging and decision making. *J Digit Imaging*, 6 (2):117-125.
- Yabuuchi H, Matsuo Y, Okafuji T, Kamitani T, Soeda H, Setoguchi T, Sakai S, Hatakenaka M, Kubo M, Sadanaga N, Yamamoto H, Honda H. 2008. Enhanced mass on contrast-enhanced breast MR imaging: Lesion characterization using combination of dynamic contrast-enhanced and diffusion-weighted MR images. *J Magn Reson Imaging*, 28 (5):1157-1165.
- Yabuuchi H, Matsuo Y, Kamitani T, Setoguchi T, Okafuji T, Soeda H, Sakai S, Hatakenaka M, Kubo M, Tokunaga E, Yamamoto H, Honda H. 2009. Non-mass-like enhancement on contrast-enhanced breast MR imaging: Lesion characterization using combination of dynamic contrast-enhanced and diffusion-weighted MR images. *Eur J Radiol*.
- Yang F. 2006. Korrelation des Ödemzeichens bei Mammaläsionen in der MR-Mammographie [Dissertation]. Jena: Friedrich-Schiller-Universität.
- Youk JH, Kim EK, Kim MJ, Oh KK. 2008. Sonographically guided 14-gauge core needle biopsy of breast masses: a review of 2,420 cases with long-term follow-up. *AJR Am J Roentgenol*, 190 (1):202-207.
- Zwick S, Brix G, Tofts PS, Strecker R, Kopp-Schneider A, Laue H, Semmler W, Kiessling F. 2009. Simulation-based comparison of two approaches frequently used for dynamic contrast-enhanced MRI. *Eur Radiol*.

9 Anhang

9.1 Beispiel und Formelherleitungen

9.1.1 Berechnung der Malignitätswahrscheinlichkeit

Einer in der MRM auffälligen pathologischen Anreicherung konnten folgende Merkmalsausprägungen zugeordnet werden: *inhomogenes Anreicherungsmuster*, im Vergleich zum umgebenden Mammaparenchym in der T2-Wichtung *hypointens*, *Root-Zeichen positiv*, *kein Ödem*, *wash-out*. Es soll die Wahrscheinlichkeit nach dem in dieser Arbeit entwickelten Fünf-Deskriptorenmodell berechnet werden.

Die Wahrscheinlichkeit für das Vorliegen eines Malignoms errechnet sich mit

$$P = \frac{1}{1 + e^{-z}}$$

wobei $z = \text{Konstante} + \text{Koeffizient}_1 * 1 + \text{Koeffizient}_2 * 1 + \dots + \text{Koeffizient}_5 * 1$
(Multiplikation eines Koeffizienten mit 1, wenn Koeffizient zutrifft)

Die Koeffizientenwerte sind der Tab. 9.1 zu entnehmen.

Für o.g. Fall errechnet sich P wie folgt:

$$z = -0,997 + 1 * 1,426 + 1 * 2,275 + 0,971 + 1 * 0 + 1 * 0 = 3,675$$

$$P = 1 / (1 + e^{-3,675})$$

$P = 97,52\%$ (dies entspräche nach den Maßstäben dieser Arbeit einem Befund, der aller höchster Wahrscheinlichkeit nach maligne ist)

Tabelle 9.1: Deskriptorvariablen für die Berechnung der Malignitätswahrscheinlichkeit

Deskriptorvariable	Ausprägung	Koeffizient für Wahrscheinlichkeits- berechnung
Binnenstruktur	homogen	0
	zentrifugal	-0,090
	inhomogen	0,971
	zentripetal	1,757
T2-Helligkeit	isointens/hypointens	0
	hyperintens	-1,571
Root-Zeichen	negativ	0
	positiv	1,426
Ödem	negativ	0
	diffus ipsilateral	0,709
	perifokal	1,238
Kurvenverlauf	persistierend	0
	Plateau	1,425
	wash-out	2,275

Konstante = -0,997

9.1.2 Formelherleitung

Nachfolgend soll die Formelherleitung zum Auffinden von Merkmalskonstellationen für mind. 95%ige Malignitätswahrscheinlichkeit dargestellt werden.

Die Wahrscheinlichkeit für das Vorliegen eines Malignoms errechnet sich mit

$$P = \frac{1}{1 + e^{-z}}$$

Wobei $z = \text{Konstante} + \text{Koeffizient}_1 * 1 + \text{Koeffizient}_2 * 1 + \dots + \text{Koeffizient}_n * 1$

(Multiplikation eines Koeffizienten mit 1, wenn Koeffizient zutrifft)

Errechnung des z-Grenzwertes, für den $P \geq 95\%$ wird. Die Formelherleitung erfordert mehrere termische Umformungen und die Anwendung logarithmischer Funktionen.

$$0,95 \leq \frac{1}{1 + e^{-z}}$$

$$0,95 * (1 + e^{-z}) \leq 1$$

$$1 + e^{-z} \leq \frac{1}{0,95}$$

$$e^{-z} \leq \frac{1}{0,95} - 1$$

$$\ln e^{-z} \leq \ln \left(\frac{1}{0,95} - 1 \right)$$

$$-z * \ln e \leq \ln \left(\frac{1}{0,95} - 1 \right)$$

$$-z * 1 \leq \ln \left(\frac{1}{0,95} - 1 \right)$$

$$z \geq -1 * \ln \left(\frac{1}{0,95} - 1 \right)$$

➔ Ist $z \geq 2,944$ (gerundet), so ist die errechnete Malignitätswahrscheinlichkeit $\geq 95\%$.

Analog kann die Herleitung zur Errechnung von z erfolgen, wenn man eine Wahrscheinlichkeit $\leq 5\%$ erhalten will:

$$z \leq -1 * \ln\left(\frac{1}{0,05} - 1\right) \rightarrow z \leq -2,944$$

9.2 Danksagung

An dieser Stelle gebührt Herrn Prof. Dr. med. Dipl.-Chem. W. A. Kaiser mein außerordentlicher Dank für die freundliche Überlassung des Themas und sein außergewöhnliches Engagement für Forschung und Lehre. Ihm gelang es durch langjährige wissenschaftliche Tätigkeit auf dem Gebiet der Brustkrebsdiagnostik, insbesondere der MR-Mammographie, die Brustkrebserkennung stetig weiter zu verbessern, um diese auf ein höchstmögliches Niveau zu heben.

Auch Herrn Dr. med. P. A. T. Baltzer ist für die hervorragende und stets sehr kooperative Zusammenarbeit bei der Datenauswertung und Manuskripterstellung zu danken. Er ist einer der mittlerweile zahlreichen Mitstreiter von Herrn Professor Kaiser, die vom Forschervirus infiziert wurden und stets bemüht sind, diesen an die künftige Medizinergeneration weiterzugeben.

Frau Dipl.-Math. A. Brandstädt sei Dank gesagt für ihre Unterstützung bei statistischen Fragestellungen.

Mit Herrn Prof. Dr. med. Dipl.-Chem. W. A. Kaiser, Herrn Dr. med. P. Baltzer, Herrn Dr. med. M. Dietzel, Frau Dr. med. C. Deicke, Frau Dr. med. K. Ferch, Frau Dr. med. D. Fischer, Herrn Dr. med. M. Freesmeyer, Frau Dr. med. G. Gust, Frau Dr. med. A. Herzog, Herrn PD Dr. med. habil. A. Malich, Frau Dr. med. S. Marx, Frau Dr. med. C. Reuchsel, Frau Dr. med. V. Schlotzer, Frau Dr. med. A. Simon und Frau Dr. med. S. Wurdinger sollen hier sämtliche Personen gewürdigt werden, die durch ihre Mitarbeit bei der Datenevaluation in den Nachschichten maßgeblich an dem Zustandekommen des unermesslichen Datenmaterials und zuletzt auch an der Möglichkeit der Bearbeitung des Themas der vorliegenden Arbeit Anteil haben.

Von ganzen Herzen sei meinen Eltern für deren finanzielle Unterstützung sowie zahlreichen Freunden für deren emotionalen Beistand gedankt.

9.3 Curriculum vitae

Name:	Tobias Gröschel
Anschrift:	Im Planer 4, 07745 Jena
Staatsangehörigkeit:	Deutsch
Geburtsdatum:	13.08.1982
Geburtsort:	Jena
1989 – 1993	Grundschule J.-G.-Fichte in Jena
1993 – 2001	Angergymnasium in Jena
2001	Erlangen der Allgemeinen Hochschulreife (Abitur), Jena
2001 – 2002	Zivildienst am Klinikum der FSU Jena
2002 – 2003	Studium der Wirtschaftsinformatik an der FSU Jena
seit 2003	Studium der Humanmedizin an der FSU Jena
03/2006	Erster Abschnitt der Ärztlichen Prüfung
08/2010 – 11/2010	Aufenthalt in Buenos Aires (Argentinien)
06/2011	Zweiter Abschnitt der Ärztlichen Prüfung
07/2011	Erwerb der Approbation als Arzt

Ort, Datum

Unterschrift

9.4 Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass mir die Promotionsordnung der Medizinischen Fakultät der Friedrich-Schiller-Universität bekannt ist,

ich die Dissertation selbst angefertigt habe und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben sind,

mich folgende Personen bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts unterstützt haben: Herr Prof. Dr. med. Dipl.-Chem. W. A. Kaiser, Herr Dr. med. P. Baltzer, Herr Dr. med. M. Dietzel, Frau Dr. med. C. Deicke, Frau Dr. med. K. Ferch, Frau Dr. med. D. Fischer, Herr Dr. med. M. Freesmeyer, Frau Dr. med. G. Gust, Frau Dr. med. A. Herzog, Herr PD Dr. med. habil. A. Malich, Frau Dr. med. S. Marx, Frau Dr. med. C. Reuchsel, Frau Dr. med. V. Schlotzer, Frau Dr. med. A. Simon, Frau Dr. med. S. Wurdinger, Frau Dipl.-Math. Antje Brandstädt,

die Hilfe eines Promotionsberaters nicht in Anspruch genommen wurde und dass Dritte unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen,

dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe und

dass ich die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung nicht bei einer anderen Hochschule als Dissertation eingereicht habe.

Ort, Datum

Unterschrift des Verfassers

